Development and validation of a multiple factor-based prognostic score system of thrombosis in polycythemia vera (MFPS-PV)

Wenjing Gu
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Yuhui Zhang
The Second Affiliated Hospital of Tianjin Medical University

Ting Sun
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Ju Mankai
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Liu Xiaofan
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Feng Xue
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Yunfei Chen
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Liu Wei
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Donglei Zhang
Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Li Huiyuan
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Wentian Wang
Institute of Hematology and Blood Disease Hospital
Ying Chi
Institute of Hematology and Blood Disease Hospital

Renchi Yang
Institute of Hematology and Blood Disease Hospital

Jie Bai
Chinese Academy of Medical Sciences and Peking Union Medical College, Institute of Hematology and Blood Diseases Hospital

Rongfeng Fu
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

Lei Zhang (✉ zanglei1@ihcams.ac.cn)
Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College

---

**Article**

**Keywords:**

**Posted Date:** May 13th, 2022

**DOI:** https://doi.org/10.21203/rs.3.rs-1630456/v1

**License:** 🎨 This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Thrombosis is an important cause of death in patients with polycythemia vera (PV). This study aimed to develop and validate a multiple factor-based prediction model of thrombosis and to propose risk-adapted treatment strategies for the 2016 World Health Organization-defined PV. The study involved 301 patients in the training cohort and another 194 patients in the external validation cohort. Multivariate analysis indicated that age ≥ 57 years (hazard ratio [HR] 2.586, p = 0.006), cardiovascular risk factors (HR 4.599, p = 0.005), previous thrombosis (HR 4.780, p < 0.001), and at least one high-risk mutation for thrombosis (mutations in DNMT3A, ASXL1, and BCOR/BCORL1) (HR 3.732, p < 0.001) were independent risk factors for thrombosis. After assigning HR-weighted scores to each risk factor, a multiple factor-based prognostic score system of thrombosis (MFPS-PV) was developed, classifying patients into low-risk (0–1 points), intermediate-risk (2–3 points), and high-risk (≥ 4 points) groups. Patients in the three groups had significantly different thrombosis-free survival rates (p < 0.001). The MFPS-PV outperformed the conventional two-tiered stratification (C-index 0.746 vs 0.643) and remained consistent during external validation. Finally, a risk-adapted therapeutic strategy was established based on the MFPS-PV. In conclusion, the MFPS-PV, integrating genetic and clinical characteristics for the first time, can significantly predict thrombosis and has great therapeutic implications.

Introduction

Polycythemia vera (PV) is a subtype of classical Philadelphia chromosome-negative (Ph−) myeloproliferative neoplasms (MPNs), characterized by an elevated red blood cell mass. Although patients with PV have relatively long survival (median 15 years) and low incidence of myelofibrosis or leukemia transformation (4.9–6% and 2.3–14.4% at 10 years, respectively), some of them have a shorter life expectancy than the age- and sex-matched general populations.[1, 2, 3, 4] Thrombosis is an important cause of death and decline in patients’ quality of life. In a study of 665 patients with PV from the Mayo Clinic, the incidence of thrombosis was 25% before diagnosis and 24% after diagnosis.[4] The large-cohorts of the European Collaboration on Low-Dose Aspirin in Polycythemia Vera (ECLAP) and International Working Group for MPN Research and Treatment (IWG-MRT) studies have reported a cumulative rate of 5.5 and 2.62 per 100 person-years of thrombotic events, respectively, in patients with PV, and mortality due to thrombosis contributed to 45% of all-cause deaths in the ECLAP trial.[5, 6] To reduce the incidence of life-threatening thrombosis in PV, it’s necessary to accurately identify thrombotic risk factors and to make appropriate antithrombotic prophylaxis.

The main goal of treatment for patients with PV is to prevent thrombotic complications. Currently, the recommended treatment strategy for PV is based on a conventional two-tiered risk stratification that classifies patients into two risk groups: “low risk,” age < 60 years and no previous thrombosis and “high risk,” age ≥ 60 years and/or with a previous thrombosis.[1, 7, 8] However, conventional risk stratification is based on the Polycythemia Vera Study Group (PVSG) criteria.[9, 10] Additionally, conventional risk stratification is specified when next-generation sequencing (NGS) technology is not widely used in clinical practice, and in fact, recent studies have highlighted the role of clonal hematopoiesis and specific
mutations in thrombosis of MPNs.

Moreover, additional risk factors for thrombosis, such as cardiovascular risk factors (CVF), higher JAK2 V617F allele burden, leukocytosis, elevated hematocrit (HCT) levels, and uncontrolled thrombocytosis have also been identified. Thus, the two-tiered stratification may ignore some potential risk factors of thrombosis in the 2016 World Health organization (WHO)-defined PV.

In this study, we integrated genetic and clinical information into thrombosis risk stratification and aimed to improve the prediction power of thrombosis in the 2016 WHO-defined PV and propose risk-adapted treatment strategies according to the new model.

**Subjects And Methods**

**Patients and samples**

The study was approved by the Institutional Ethics Committee of the Institute of Hematology and Blood Diseases Hospital (Tianjin, China) and the Second Hospital of Tianjin Medical University (Tianjin, China). Patients at the two institutions were eligible for inclusion if 18 years or older with a confirmed diagnosis of PV according to the 2016 WHO criteria. Clinical and laboratory data at the initial diagnosis were retrospectively collected from institutional databases, and additional information during follow-up was obtained by direct contact with patients or their families. Bone marrow mononuclear cells were collected from the included patients and sent for NGS. Data from the Institute of Hematology and Blood Diseases Hospital were initially analyzed for model development, while the cohort from the Second Hospital of Tianjin Medical University was subsequently used for external validation.

**Next-generation sequencing**

NGS was performed on 1 µg of genomic DNA extracted from the patients’ bone marrow mononuclear cells. Sequencing was performed using the NovaSeq 6000 platform (Illumina, San Diego, CA, USA) with a targeted 267-gene panel. A list of 267 candidate genes is provided in Supplemental Table 1. After filtering out low-quality sequencing reads, clean reads were mapped to the human reference genome (hg19) using the Burrows–Wheeler Aligner. Single-nucleotide polymorphisms and short insertion/deletions were called and annotated using 1000 Genomes, ESP6500, SNPeffect, Inhouse, PolyPhen-2, SIFT, and COSMIC to determine the potential functional impacts of the variants. The identified pathogenic mutations were validated by Sanger sequencing. Details of the sequencing and bioinformatics analysis are provided in the Supplemental Methods.

**Statistical analysis**

Categorical variables were compared using the chi-square test, while continuous variables were compared using the Mann–Whitney U test or Kruskal–Wallis test. Receiver operating characteristic plots were used to define cutoff levels for continuous variables that were significant in the univariate analysis. A multivariate Cox proportional hazard regression analysis was conducted to identify independent risk factors for thrombosis-free survival (TFS). Hazard ratio (HR) with 95% confidence intervals (CI) and
corresponding p-values were calculated. HR-based risk score allocation was employed to obtain a multifactorial prediction model, as previously reported.[24] Survival curves were constructed using the Kaplan–Meier method and compared using the log-rank test. TFS was considered from the date of diagnosis to the date of the first episode of thrombosis after diagnosis (uncensored) or the last contact (censored). Discrimination power was assessed using the Harrell C-concordance index, where a C-index of 1.00 indicates perfect discrimination, while a value of 0.5 indicates random discrimination. The prediction model was considered to have good discriminative capability if the concordance index was > 0.7.[25] Analyses were performed using the SPSS software (version 26.0), and graphs were constructed using the R software (version 4.0.2). Statistical significance was set at p < 0.05.

Results

Patient characteristics and incidence of thrombosis

All data were censored on September 30 2021. A total of 301 NGS-annotated PV patients from the Institute of Hematology and Blood Diseases Hospital were included in the training cohort, with a median follow-up of 43 months (range 12–454). The clinical and laboratory characteristics of patients at diagnosis are summarized in Table 1. The median age at diagnosis was 54 years (range 18–86) and 58.5% were male. Of the patients, 201 (66.8%) had at least one CVF (including hypertension, diabetes, hyperlipidemia, obesity, and smoking). All patients harbored clearly elevated HCT levels (median 58.2%; range 46.4–74.3) or hemoglobin (median 187 g/L; range 160–290) concentration, and most of them (n = 224/268, 83.6%) had suppressed erythropoietin concentration (median 1.35 IU/L; range 0.19–17.73). Abnormal karyotype and palpable splenomegaly were documented in 11.5% and 81.2% of the patients, respectively.
Table 1
Comparison of clinical and mutation information of 301 PV patients with or without thrombosis after diagnosis.

<table>
<thead>
<tr>
<th></th>
<th>All patients (n = 301)</th>
<th>Patients with thrombosis after diagnosis (n = 50)</th>
<th>Patients without thrombosis after diagnosis (n = 251)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, median (range), y</td>
<td>54 (18–86)</td>
<td>61 (21–81)</td>
<td>52 (18–86)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Gender, male, n (%)</td>
<td>176 (58.5)</td>
<td>28 (56.0)</td>
<td>148 (59.0)</td>
<td>0.698</td>
</tr>
<tr>
<td>CVF*, n (%)</td>
<td>201 (66.8)</td>
<td>46 (92.0)</td>
<td>155 (61.8)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Previous thrombosis, n (%)</td>
<td>95 (31.6)</td>
<td>34 (68.0)</td>
<td>61 (24.3)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>Laboratory characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemoglobin, median (range), g/L</td>
<td>187 (160–290)</td>
<td>192 (163–228)</td>
<td>186 (160–290)</td>
<td>0.227</td>
</tr>
<tr>
<td>Hematocrit, median (range), %</td>
<td>58.2 (46.4–74.3)</td>
<td>59.0 (47.8–73.4)</td>
<td>57.2 (46.4–74.3)</td>
<td>0.411</td>
</tr>
<tr>
<td>Leukocytes, median (range), ×10⁹/L</td>
<td>11.8 (4.3–50.0)</td>
<td>11.7 (6.6–33.0)</td>
<td>11.8 (2.5–50.0)</td>
<td>0.635</td>
</tr>
<tr>
<td>Platelets, median (range), ×10⁹/L</td>
<td>494 (109–1609)</td>
<td>471 (115–1241)</td>
<td>506 (100–1609)</td>
<td>0.474</td>
</tr>
<tr>
<td>Erythropoietin, (IU/L)</td>
<td>1.35 (0.19–17.73)</td>
<td>1.50 (0.48–11.71)</td>
<td>1.33 (0.19–17.73)</td>
<td>0.563</td>
</tr>
<tr>
<td>Palpable splenomegaly, n (%)</td>
<td>194/239 (81.2)</td>
<td>33/42 (76.2)</td>
<td>161/197 (81.7)</td>
<td>0.635</td>
</tr>
<tr>
<td>Abnormal karyotype, n (%)</td>
<td>33/287 (11.5)</td>
<td>7/48 (14.6)</td>
<td>26/239 (10.9)</td>
<td>0.463</td>
</tr>
<tr>
<td><strong>Driver mutations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JAK2 V617F</td>
<td>279 (92.7)</td>
<td>49 (98.0)</td>
<td>230 (91.6)</td>
<td>0.200</td>
</tr>
</tbody>
</table>

*CVF, cardiovascular risk factors, including hypertension, diabetes, hyperlipidemia, obesity or smoking.
#Genes with mutations in at least two patients in the training cohort are displayed and compared by χ² test. PV, polycythemia vera; VAF, variant allele frequency. Bold indicates statistically significant values.
<table>
<thead>
<tr>
<th>Mutation</th>
<th>All patients (n = 301)</th>
<th>Patients with thrombosis after diagnosis (n = 50)</th>
<th>Patients without thrombosis after diagnosis (n = 251)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAK2 exon12</td>
<td>9 (3.0)</td>
<td>1 (2.0)</td>
<td>8 (3.2)</td>
<td>1.000</td>
</tr>
<tr>
<td>JAK2 negative</td>
<td>13 (4.3)</td>
<td>0</td>
<td>13 (5.2)</td>
<td>0.206</td>
</tr>
<tr>
<td>JAK2 V617F</td>
<td>51.0 (1.1–96.8)</td>
<td>59.0 (6.1–95.8)</td>
<td>49.7 (1.1–96.8)</td>
<td>0.072</td>
</tr>
<tr>
<td>JAK2 exon12</td>
<td>20.5 (10.0–82.2)</td>
<td>45.7</td>
<td>20.3 (10.0–82.2)</td>
<td>0.245</td>
</tr>
<tr>
<td>Other mutations#</td>
<td>71 (23.6)</td>
<td>23 (46.0)</td>
<td>48 (19.1)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>TET2</td>
<td>34 (11.3)</td>
<td>7 (14.0)</td>
<td>27 (10.8)</td>
<td>0.508</td>
</tr>
<tr>
<td>ASXL1</td>
<td>24 (8.0)</td>
<td>10 (20.0)</td>
<td>14 (5.6)</td>
<td>0.002</td>
</tr>
<tr>
<td>DNMT3A</td>
<td>23 (7.6)</td>
<td>11 (22.0)</td>
<td>12 (4.8)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>BCOR/BCORL1</td>
<td>21 (7.0)</td>
<td>8 (16.0)</td>
<td>13 (5.2)</td>
<td>0.015</td>
</tr>
<tr>
<td>TP53</td>
<td>6 (2.0)</td>
<td>1 (2.0)</td>
<td>5 (2.0)</td>
<td>1.000</td>
</tr>
<tr>
<td>SF3B1</td>
<td>4 (1.3)</td>
<td>0</td>
<td>4 (1.6)</td>
<td>0.824</td>
</tr>
<tr>
<td>CBL</td>
<td>2 (0.7)</td>
<td>0</td>
<td>2 (0.8)</td>
<td>1.000</td>
</tr>
<tr>
<td>SETBP1</td>
<td>2 (0.7)</td>
<td>0</td>
<td>2 (0.8)</td>
<td>1.000</td>
</tr>
<tr>
<td>Epigenetic regulation</td>
<td>71 (23.6)</td>
<td>23 (46.0)</td>
<td>48 (19.1)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>mRNA splicing</td>
<td>5 (1.7)</td>
<td>0</td>
<td>5 (2.0)</td>
<td>0.689</td>
</tr>
<tr>
<td>Signaling pathways</td>
<td>6 (2.0)</td>
<td>2 (4.0)</td>
<td>4 (1.6)</td>
<td>0.577</td>
</tr>
<tr>
<td>Transcription regulation</td>
<td>3 (1.0)</td>
<td>0</td>
<td>3 (1.2)</td>
<td>1.000</td>
</tr>
<tr>
<td>Cell cycle/apoptosis</td>
<td>26 (8.6)</td>
<td>9 (18.0)</td>
<td>17 (6.8)</td>
<td>0.078</td>
</tr>
</tbody>
</table>

*CVF, cardiovascular risk factors, including hypertension, diabetes, hyperlipidemia, obesity or smoking. 
Genes with mutations in at least two patients in the training cohort are displayed and compared by χ² test. PV, polycythemia vera; VAF, variant allele frequency. Bold indicates statistically significant values.


<table>
<thead>
<tr>
<th></th>
<th>All patients (n = 301)</th>
<th>Patients with thrombosis after diagnosis (n = 50)</th>
<th>Patients without thrombosis after diagnosis (n = 251)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Follow-up, median (range), months</strong></td>
<td>43 (12–454)</td>
<td>55 (12–304)</td>
<td>42 (12–454)</td>
<td>0.076</td>
</tr>
</tbody>
</table>

*CVF, cardiovascular risk factors, including hypertension, diabetes, hyperlipidemia, obesity or smoking. #Genes with mutations in at least two patients in the training cohort are displayed and compared by χ² test. PV, polycythemia vera; VAF, variant allele frequency. Bold indicates statistically significant values.

Altogether, 111 patients (36.9%) suffered from thrombosis, including 105 (34.9%) with arterial thrombosis and 12 (4.0%) with venous thrombosis. Regarding the occurrence of thrombosis, 48 patients (15.9%) had thrombosis before diagnosis, 59 patients (19.6%) at diagnosis, and 50 patients (16.6%) after diagnosis (Supplemental Table 2).

**Mutation landscapes and identification of high-risk mutations for thrombosis**

The driver mutation distribution was JAK2 V617F in 279 (92.7%) and JAK2 exon12 in nine (3.0%) patients. Thirteen (4.3%) patients had no driver mutations. Mutations other than JAK2 V617F or JAK2 exon12 were detected in 103 (34.2%) patients, with the most frequently mutated genes being involved in epigenetic regulation, such as TET2, ASXL1, and DNMT3A (Table 1, Fig. 1, and Supplemental Table 3). No association was observed between the risk of thrombosis and JAK2 V617F (p = 0.200) or JAK2 exon12 mutations (p = 1.000), whereas mutations in genes encoding for epigenetic modifiers were more frequent in patients with thrombosis than in those without thrombosis after diagnosis (46.0% vs. 19.1%, p < 0.001). Specifically, significant differences were noted in the frequencies of ASXL1 (20.0% vs. 5.6%, p = 0.002), DNMT3A (22.0% vs. 4.8%, p < 0.001) and BCOR/BCORL1 (16.0% vs. 5.2%, p = 0.015) mutations between patients with and without thrombosis after diagnosis. The variant allele frequency of JAK2 V617F and JAK2 exon12 mutations was also analyzed, and no significant difference was identified (p = 0.072 and p = 0.245, respectively) (Table 1).

**Risk factors for thrombosis after diagnosis**

Demographic, clinical, cytogenetic, and molecular genetic features were included in the univariate analysis (Table 1). In the univariate analysis, the risk factors for thrombosis after diagnosis were age at diagnosis (p < 0.001), CVF (p < 0.001), previous thrombosis (p < 0.001), epigenetic mutations (p < 0.001), and specific high-risk mutations for thrombosis (ASXL1, DNMT3A, and BCOR/BCORL1, p < 0.05 for each mutation). A receiver operating characteristic plot was used to define the optimal cutoff level for age, and the best discriminant value was 57 years (Supplemental Fig. 1). The TFS curves are displayed in Supplemental Fig. 2, demonstrating the significance of each risk factor (p < 0.001). In the multivariable
Cox proportional hazard regression analysis, we identified four independent risk factors of thrombosis after diagnosis, including age \( \geq 57 \) years (HR 2.586, 95%CI, 1.321–5.063, \( p = 0.006 \)), CVF (HR 4.599, 95%CI, 1.573–13.446, \( p = 0.005 \)), previous thrombosis (HR 4.780, 95%CI, 2.455–9.308, \( p < 0.001 \)), and at least one high-risk mutation for thrombosis (HR 3.732, 95%CI, 2.121–6.566, \( p < 0.001 \)) (Table 2).

### Table 2
Multivariable Cox regression analysis of risk factors that are predictive of thrombosis after diagnosis.

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>HR</th>
<th>95%CI</th>
<th>( P )</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age ( \geq 57 ) years</td>
<td>2.586</td>
<td>1.321–5.063</td>
<td>0.006</td>
<td>1</td>
</tr>
<tr>
<td>CVF</td>
<td>4.599</td>
<td>1.573–13.446</td>
<td>0.005</td>
<td>2</td>
</tr>
<tr>
<td>Previous thrombosis</td>
<td>4.780</td>
<td>2.455–9.308</td>
<td>&lt; 0.001</td>
<td>2</td>
</tr>
<tr>
<td>High-risk mutation for thrombosis*</td>
<td>3.732</td>
<td>2.121–6.566</td>
<td>&lt; 0.001</td>
<td>1</td>
</tr>
</tbody>
</table>

*At least one high-risk mutation for thrombosis, including ASXL1, DNMT3A and BCOR/BCORL1. HR, hazard ratio; CI, confidence interval; CVF, cardiovascular risk factor.

### Development of a multiple factor-based prognostic score system of thrombosis in PV (MFPS-PV)

To develop a multiple factor-based thrombosis scoring system, HR-weighted risk points were assigned to each risk factor identified in the multivariate analysis (Table 2, age \( \geq 57 \) years, 1 point; CVF, 2 points; previous thrombosis, 2 points; at least one high-risk mutation for thrombosis, 1 point). After summarizing the points, patients were divided into low-risk (\( n = 85, 28.2\% \), score \( 0–1 \)), intermediate-risk (\( n = 118, 39.2\% \), score \( 2–3 \)), and high-risk (\( n = 98, 32.6\% \), score \( \geq 4 \)) groups. The incidence of thrombosis after diagnosis in the low-, intermediate-, and high-risk groups was 0.32, 1.49, and 12.93/100 person-years, with an estimated 10-year TFS of 95.6\%, 83.7\%, and 15.2\%, respectively (Table 3). Significantly different TFS curves were observed among the three risk groups by Kaplan–Meier analysis (\( p < 0.001 \)) (Fig. 2A).

### Table 3
Distribution of scores and risk stratification among 301 PV patients based on the MFPS-PV.

<table>
<thead>
<tr>
<th>Score</th>
<th>Risk stratification (n, %)</th>
<th>Incidence of thrombosis (/100 person-years)</th>
<th>Ten-year TFS, % (95%CI)</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–1</td>
<td>Low risk (85, 28.2)</td>
<td>0.32</td>
<td>95.6 (89.5–100.0)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>2–3</td>
<td>Intermediate risk (118, 39.2)</td>
<td>1.49</td>
<td>83.7 (70.0–100.0)</td>
<td></td>
</tr>
<tr>
<td>( \geq 4 )</td>
<td>High risk (98, 32.6)</td>
<td>12.93</td>
<td>15.2 (4.7–49.4)</td>
<td></td>
</tr>
</tbody>
</table>

PV, polycythemia vera; TFS, thrombosis-free survival; CI, confidence interval.
Comparison of discrimination power between the MFPS-PV and conventional stratification

To verify the discrimination power of the MFPS-PV, the risk classification of 301 patients with PV according to the MFPS-PV or conventional stratification was compared. As shown in Fig. 2B, 73 patients (47.4%) in the conventional low-risk group were still confirmed to be at low risk by MFPS-PV, whereas 77 patients (50.0%) were regrouped as intermediate-risk, and four patients (2.6%) were classified as high-risk. Among the four patients who were regrouped as high-risk, the actual incidence of thrombosis was high at 6.1/100 person-years. Additionally, 94 patients (63.9%) in the conventional high-risk group were also considered to be at high risk according to the MFPS-PV, while the remaining 53 patients (36.1%) were regarded as the lower risk category. Among the 53 regrouped patients, only four experienced thrombosis during follow-up, with an incidence of thrombosis of 2.38/100 person-years. Therefore, MFPS-PV can identify high-risk patients ignored by conventional stratification and non-high-risk patients over-evaluated by conventional stratification, so as to avoid insufficient or excessive treatment of these patients.

Based on the Harrell C-statistical analysis, the C-index was 0.711 for conventional stratification and 0.794 for the MFPS-PV, supporting the better discrimination power of MFPS-PV (Supplemental Fig. 3). Therefore, MFPS-PV better predicted the risk of thrombosis than conventional two-tiered stratification.

External validation of MFPS-PV

To validate the predictive accuracy of the MFPS-PV, 194 WHO (2016)-defined PV patients from the Second Hospital of Tianjin Medical University were assessed as an external validation cohort. The median age was 55 years (range 21–85) at diagnosis and the sex distribution was close to 1:1 (male, 49.0%) (Supplemental Table 4). During a median follow-up of 96 months (range 12–576), 91 (47.0%) patients experienced thrombosis after diagnosis. The MFPS-PV model applied to the external cohort demonstrated excellent discrimination in low-risk (n = 60), intermediate-risk (n = 77), and high-risk (n = 57) patients, with corresponding 10-year TFS rates of 81.9%, 62.8%, and 18.3%, respectively (Supplemental Table 5). TFS curves confirmed the significant difference among the three risk groups in the external cohort (n = 194, p < 0.001) and both two cohorts (n = 495, p < 0.001) of patients with PV, suggesting that the MFPS-PV has good discrimination ability in the external validation set with similar basic characteristics but much longer follow-up than the training cohort (Fig. 3).

Treatment strategies and recommendations according to the MFPS-PV

We analyzed the treatment patterns of 301 patients with PV to determine the optimal treatment regimen applicable for patients in each thrombotic risk group according to the MFPS-PV. Among the 301 patients, 153 (50.8%) received both cytoreductive and antiplatelet therapies (combination therapy), 72 (23.9%) received antiplatelet therapy only (antiplatelet monotherapy), 34 (11.3%) received cytoreductive therapy only (cytoreductive monotherapy), and 42 (14.0%) did not receive any drug therapy. In addition to drug
administration, 82 (27.2%) patients underwent regular therapeutic phlebotomy (TP). Patients receiving TP accounted for 29.4%, 32.4%, 26.4%, and 16.7% of the patients receiving combination therapy, cytoreductive monotherapy, antiplatelet monotherapy, and no therapy, respectively. Details of treatments for patients in each risk group were displayed in Fig. 4A and Supplemental Table 6.

For patients with PV, the antithrombotic value of low-dose aspirin and TP has been confirmed in controlled studies.[26, 27] Therefore, it is critical to identify the appropriate patients for cytoreductive therapy. We compared the incidence of thrombosis according to different treatment regimens in low-, intermediate-, and high-risk patients, respectively. In the high-risk group, patients receiving antiplatelet monotherapy had a lower thrombosis rate than those without drug therapy (30.3% vs. 94.1%, p < 0.001), whereas patients receiving combination therapy had a lower thrombosis rate than both patients receiving antiplatelet monotherapy (10.8% vs. 30.3%, p = 0.042) and those without drug therapy (10.8% vs. 94.1%, p < 0.001). The results suggest that high-risk patients benefited more from combination therapy. In the low-risk group, no difference in thrombosis rates was observed among patients who received combination therapy, antiplatelet monotherapy, or those without drug therapy (p > 0.05 for each comparison). In the intermediate-risk group, we observed lower thrombosis rates in patients who received combination therapy (3.1% vs. 28.6%, p = 0.007) and antiplatelet monotherapy (0 vs. 28.6%, p = 0.014) than in those without drug therapy. However, no significant differences in the incidence of thrombosis were identified between the combination therapy and antiplatelet monotherapy groups (p = 1.000). Therefore, adding cytoreduction may not be necessary in intermediate-risk patients. By further division, we discovered that intermediate-risk patients with two thrombotic risk factors benefited more from combination therapy than from no drug therapy (p = 0.005) (Fig. 4B). Finally, we analyzed the impact of cytoreductive treatment on patients with different risk factors. Patients with previous thrombosis or CVF treated with combination therapy had a lower risk of thrombosis than those treated with antiplatelet monotherapy (p = 0.022 and p = 0.011, respectively), suggesting that adding cytoreduction was necessary in intermediate-risk patients with previous thrombosis or CVF (Fig. 4C).

To improve thromboprophylaxis, we proposed our risk-adapted therapy recommendations for patients with PV based on MFPS-PV (Fig. 5). TP and low-dose antiplatelet therapy are recommended for all patients to achieve a targeted HCT level lower than 45%. Low-risk (0–1 score) patients are not mandatory for cytoreduction and only require low-dose antiplatelet therapy. Intermediate-risk (2–3 scores) patients are advised for antiplatelet therapy, with cytoreduction added if they have two thrombotic risk factors, previous thrombosis, or CVF. High-risk (≥ 4 scores) patients are recommended antiplatelet and cytoreductive combination therapy.

**Discussion**

In this study, we analyzed clinical information and NGS data from a large cohort of patients with PV to reveal the risk factors for thrombosis. After comprehensive analysis, we identified age ≥ 57 years, CVF, previous thrombosis, and high-risk mutations (e.g., DNMT3A, ASXL1, and BCOR/BCORL1 mutations) as risk factors for thrombosis after diagnosis and then established a multiple factor-based prediction model.
classifying patients into low-risk, intermediate-risk, and high-risk groups. This new model incorporates genetic information into thrombosis prediction and put emphasis on development of precise thromboprophylaxis strategies applicable to patients in different risk groups.

Advanced age and prior thrombosis have been recognized as risk factors in the conventional two-tiered stratification since the publication of ECLAP trial which included 1638 patients.\[^5\] Compared with conventional stratification, the new model has some differences and potential advantages.

Firstly, the MFPS-PV model was developed based on 2016 WHO-defined PV\[^23\] whereas the conventional model is based on PVSG-defined PV.\[^9, 10\] The 2016 WHO-defined PV recognizes the importance of bone marrow morphology, thus patients previously considered to be “masked” or “prodromic” PV were included according to the new criteria.\[^28, 29\] In fact, “masked” or “prodromic” PV displays a higher risk of thrombosis in younger patients probably due to the lower intensity of treatment.\[^30\] Hence, MFPS-PV is more suitable for patients diagnosed using the current diagnostic standard than the traditional model.

Secondly, the MFPS-PV model highlights the significant predictive value of CVF. The current guideline recommendations from the National Comprehensive Cancer Network clearly emphasize that CVF in patients with PV need to be controlled.\[^7\] Previous studies have also reported that CVF may increase the risk of thrombosis in patients with PV.\[^14, 15\] Moreover, some experts commonly provide different treatment suggestions according to the presence or absence of CVF, as it plays an important role in the prevention and treatment of thrombosis.\[^1\]

Thirdly and notably, we have incorporated genetic aberrations into the prognostic stratification system for the first time. The wide application of NGS in a large cohort of patients with PV allows us to evaluate the impact of specific mutations on thrombosis, which has not been mentioned in most previous studies. Several thrombogenic mutations have been proposed in our study, including mutations in DNMT3A, ASXL1, and BCOR/BCORL1. The epigenetic regulators DNMT3A and ASXL1 are commonly mutated in MPNs and have been proven to be associated with disease initiation and progression.\[^31–33\] Previous reports have also confirmed that mutations in genes encoding epigenetic regulators are associated with increased cardiovascular risks and atherosclerosis, mainly due to upregulation of pro-inflammatory signaling triggered by macrophage deficiency.\[^34, 35\] Moreover, a previous study has demonstrated that mutations in DNMT3A, TET2, and ASXL1 were strongly correlated with thrombosis occurrence, even if they evaluated a limited number of patients with PV.\[^11\] A recent study has also reported that mutations in TET2, DNMT3A, and ASXL1 are significantly associated with prior stroke.\[^13\] Those relevant findings partially support our conclusion. BCOR (BCL6 Corepressor) and BCORL1 (BCL6 Corepressor like 1) are two homologous genes located on chromosome X that have been described in approximately 4–6% of myelodysplastic syndrome cases and 16% of blastic phase chronic myeloid leukemia cases. Mutations in BCOR/BCORL1 are associated with shortened survival in myeloid malignancies.\[^36–39\] In aplastic anemia, patients with BCOR and BCORL1 mutations have an improved response to immunosuppression.\[^40\] Although reports of BCOR/BCORL1 mutations involved in Ph\(^+\) MPNs are rare, there is still evidence supporting their role in treatment resistance and poor prognosis.\[^41–43\] To the best of our knowledge,
this is the first study to reveal that BCOR/BCORL1 mutations may increase the risk of thrombosis in patients with PV.

Fourthly, the MFPS-PV displayed better discrimination power than conventional stratification in predicting thrombosis. When regrouping according to the MFPS-PV, nearly half of the patients in the conventional low-risk group were regrouped as intermediate- or high-risk. Meanwhile, more than one-third of the patients in the conventional high-risk group were regarded as having low or intermediate risk. The actual incidence of thrombosis at follow-up confirmed that MFPS-PV predicted thrombosis more accurately than the traditional model, and that MFPS-PV had the potential to avoided under- or over-therapy in some patients. In addition, the higher C-index of MFPS-PV further supported the fact that MFPS-PV outperformed the conventional one. The predictive accuracy of MFPS-PV was also validated in an external cohort with similar baseline characteristics as the training set but with a much longer follow-up.

In addition to the above advantages, there are some differences between the MFPS-PV and the traditional model. Firstly, the cutoff value of age (57 years) for thrombosis prediction was slightly lower than that reported in previous investigations, mainly 60 years old.[14, 44] This may be due to differences in the prevalence of thrombosis between Chinese and Caucasian populations. According to a global epidemiologic report of thrombosis, the average age at first thrombosis was lower in developing countries where people < 50 years of age frequently experience ischemic heart disease, than in developed countries. [45] Additionally, the incidence of thrombosis is elevated among younger persons (18–45 years) due to the increasing prevalence of CVF in such population.[46] In a large cohort of Chinese patients with cardiovascular disease, the prevalence of coronary heart disease and stroke increased markedly from the age of 40 years.[47] Another study including 783 patients with deep vein thrombosis has reported three frequency peaks, including two smaller peaks at ages 20–24 and 70–74 years and the largest peak at age 45–59 years.[48] In other studies conducted in China, the median age was only 52 years in patients suffering from the first venous thrombosis and even lower in patients suffering from cerebral venous sinus thrombosis (median, 37 years).[49–50] These results indicate that improved awareness of thrombosis prevention and management is urgently required in China. In this study, we present a thrombosis prediction model for PV that is suitable for the current epidemiology of thrombosis in China, and even in the East Asia. A lower age cutoff may help raise public concerns regarding early thrombosis prevention in young patients with PV.

In addition, we found that JAK2V617F mutational status or allele burden had no influence on thrombosis, which has been proven to be significant in some, but not all, studies.[16–18] The possible reasons are that JAK2V617F alterations affect thrombosis in ET rather than in PV,[51–53] and more likely, JAK2V617F mutant burden may be in salient associations with blast transformation instead of thrombosis.[54–56] Moreover, we only evaluated a single time point JAK2V617F at first diagnosis, so the significance of the changes in JAK2V617F allele burden occurring during disease evolution may be ignored. Thus, long-term sequential monitoring of JAK2V617F allele burden during follow-up should be encouraged in future studies to better predict thrombotic risks.
According to the widely accepted treatment recommendations, phlebotomy and low-dose acetyl salicylic acid are first-line therapies for all patients with PV.[1, 7, 8, 26, 27] Additionally, high-risk patients in conventional stratification are advised to receive either hydroxyurea or interferon as cytoreductive therapy. In this study, we analyzed the incidence of thrombosis after diagnosis under different treatment strategies in low-, intermediate-, and high-risk patients according to the MFPS-PV, respectively, and proposed an evidence-based risk-adapted therapeutic regime. We supported conservative management in patients in low-risk group. For these patients, antiplatelet therapy and regular phlebotomy to reach a target HCT level of <45% can effectively prevent thrombosis. Furthermore, MFPS-PV delineated thrombosis history and CVF as the two most detrimental thrombotic risk factors, and patients with previous thrombosis or CVF treated with combination therapy had a lower risk of thrombosis than those treated with antiplatelet monotherapy. Thus, cytoreduction is strongly recommended in intermediate-risk patients with one of these two risk factors. Antiplatelet and cytoreductive combination therapy is recommended for high-risk patients in order to avoid recurrent thrombosis and improve TFS.

In conclusion, by integrating basic clinical data together with thrombotic high-risk mutations, we presented a multiple factor-based, three-tiered, and four-factor thrombosis prediction model that was applicable to both training and external cohorts of the 2016 WHO-defined PV. This model improves the predictive power of thrombosis and has the potential to realize more precise risk-adapted management for patients with PV.

Declarations

Acknowledgements

This project was supported by grants from CAMS Innovation Fund for Medical Sciences (CIFMS) (2021-I2M-1-073, 2021-I2M-1-003), the National Natural Science Foundation of China (81970121, 81600099, 82000136), Tianjin Municipal Science and Technology Commission Grant (18JCQNJC11900), Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2020-PT310-011), and National Key Research and Development Program of China (2019YFA0110802).

Author Contributions

WJG collected data, analyzed data and wrote the paper. LZ, RFF and JB designed the study, analyzed data and revised the manuscript. YHZ, TS and MKJ collected data and analyzed data. XFL, FX, YFC, WL, HYL, WTW, YC and RCY helped with data collection. DLZ assisted with sequencing data processing.

Competing Interests

The authors declare no competing financial interests.

Data Availability Statement
Targeted next-generation sequencing data generated in this study are included in supplementary material. Additional information is available from the authors upon reasonable request.

References


**Figures**

**Figure 1**

*Genomic landscape of 301 PV patients with or without thrombosis after diagnosis.* Mutated genes are spread along the y-axis, and each column corresponds to a patient. Patients are grouped according to the presence (left side) or absence (right side) of thrombosis after diagnosis. Color code represents mutation status of the gene as indicated. Annotation rows at bottom include gender, karyotype, CV risk, and driver mutation. Events where information is unknown are depicted in light gray. PV, polycythemia vera; CVF, cardiovascular risk factors.
Figure 2

(A) **Thrombosis-free survival of 301 PV patients stratified by MFPS-PV.** Patients among different risk groups were compared by log-rank test.

(B) **Risk Distribution of 301 PV patients and their thrombosis rates according to the conventional two-tiered stratification and the MFPS-PV.** Colored bars represent distribution of patients according to the MFPS-PV in the context of conventional low-risk or high-risk group respectively. The incidence of thrombosis for each category is shown. Conventional stratification
for PV: High risk, age ≥ 60 years or/and thrombosis history; Low risk, absence of both factors. /100 pts-year, number of patients suffered from thrombosis after diagnosis per 100 patients-years. PV, polycythemia vera.

Figure 3

Risk-stratified thrombosis-free survival in (A) an external cohort of 194 PV patients and (B) all two cohorts of 495 PV patients. Patients among different risk groups are compared by log-rank test. PV, polycythemia vera.
Figure 4

(A) Distribution of phlebotomy and drug therapy for patients in different thrombotic risk groups. n/N, Number of patients with therapeutic phlebotomy /total number of patients in each group. (B) Incidence of thrombosis after diagnosis for patients in different thrombotic risk groups according to their treatment strategies. (C) Incidence of thrombosis after diagnosis for patients with each thrombotic risk factor according to their treatment strategies. Combination therapy, combination of both cytoreductive and antiplatelet therapy. *p<0.05, **p<0.01, ***p<0.001.
Figure 5

**Recommendations for Thrombosis Prevention in MFPS-PV**

Phlebotomy to reach a HCT level < 45%

- **Low risk (0-1 score)**: Antiplatelet therapy
- **Intermediate risk (2-3 scores)**: Antiplatelet therapy
- **High risk (≥4 scores)**: Antiplatelet + Cytoreductive therapy

*Presence of two thrombotic risk factors* or previous thrombosis or CVF

**Figure 5**

**Thrombosis prevention algorithm for polycythemia vera according to the MFPS-PV.** *Thrombotic risk factors include age≥57 years, CVF, previous thrombosis, and at least one thrombotic high-risk mutation. PV, polycythemia vera; HCT, hematocrit; CVF, cardiovascular risk factors.*

**Supplementary Files**

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

- [SupplementalMaterial.pdf](#)
- [SupplementalTable3.docx](#)