European LeukemiaNet study on the reproducibility of bone marrow features in masked polycythemia vera and differentiation from essential thrombocytemia

Hans Michael Kvasnicka1 | Attilio Orazi2 | Juergen Thiele3 | Giovanni Barosi4 | Carlos E. Bueso-Ramos5 | Alessandro M. Vannucchi6 | Robert P. Hasserjian7 | Jean-Jacques Kiladjian8 | Umberto Gianelli9 | Richard Silver10 | Tariq I. Mughal11 | Tiziano Barbui12

1Senckenberg Institute of Pathology, University of Frankfurt, Germany; 2Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, NY; 3Institute of Pathology, University of Cologne, Cologne, Germany; 4Center for the Study of Myelofibrosis, Biotechnology Research Area, Istituto di Ricovero e Cura a Carattere Scientifico Policlinico S. Matteo Foundation, Pavia, Italy; 5Department of Hematopathology, MD Anderson Cancer Center, Houston; 6Dipartimento di Medicina Sperimentale e Clinica, University of Florence, Florence, Italy; 7Department of Pathology, Massachusetts General Hospital, Boston; 8Centre d’Investigations Cliniques, Hôpital Saint-Louis, Paris, France; 9Pathology Unit, University of Milan Medical School, Fondazione IRCCS Cà Granda, Ospedale, Italy; 10Weill Cornell Medical College, New York, NY; 11Tufts Medical Center, Boston; 12Division of Hematology and Research Foundation, Ospedale Papa Giovanni XXIII, Bergamo, Italy

Correspondence
Hans Michael Kvasnicka, M.D., PhD, Senckenberg Institute of Pathology, University of Frankfurt, Theodor-Stern-Kai 8, 60590 Frankfurt, Germany. Email: hans-michael.kvasnicka@kgu.de

Funding Information
The Alpine Oncology Foundation (AOF)

Abstract
The purpose of the study was to assess consensus and interobserver agreement among an international panel of six hematopathologists regarding characterization and reproducibility of bone marrow (BM) histologic features used to diagnose early stage myeloproliferative neoplasms, in particular differentiation of so-called masked/prodromal polycythemia vera (mPV) from JAK2-mutated essential thrombocytemia (ET). The six members of the hematopathology panel evaluated 98 BM specimens independently and in a blinded fashion without knowledge of clinical data. The specimens included 48 cases of mPV according to the originally published hemoglobin (Hb) threshold values for this entity (male: 16.0–18.4 g/dL, female: 15.0–16.4 g/dL), 31 cases with overt PV according to the updated 2016 WHO criteria, and 19 control cases. The latter group included cases of JAK2-mutated ET, primary myelofibrosis, myelodysplastic syndrome, and various reactive conditions. Inter-rater agreement between the panelists was very high (overall agreement 92.6%, kappa 0.812), particularly with respect to separating mPV from ET. Virtually all cases of mPV were correctly classified as PV according to their BM morphology. In conclusion, a central blinded review of histology slides by six hematopathologists demonstrated that highly reproducible specific histological pattern characterize PV and confirmed the notion that there are no significant differences between mPV and overt PV in relation to BM morphology.

1 | INTRODUCTION

Concern has been repeatedly expressed regarding the accurate diagnosis of polycythemia vera (PV), particularly in its earlier stages.1–3 The diagnostic thresholds for this diagnosis are those proposed by the World Health Organization (WHO), principally based on hemoglobin (Hb) values4,5 and those of the British Committee for Standards in Haematology (BCSH) that are predominantly based on hematocrit (Hct) levels.6,7 Ample discussion has followed on which of the two diagnostic approaches had the highest sensitivity for a diagnosis of early PV.8,9 More recent evidence has shown that the original 2008 WHO thresholds for Hb were too high and thus not sensitive enough in diagnosing early phase PV and poorly correlated with red cell mass study.3

Hans Michael Kvasnicka, Attilio Orazi, and Juergen Thiele contributed equally to the manuscript.
Because of these data and similar observational experience, it became evident that a group of JAK2-mutated patients with PV may present at onset with a subnormal serum erythropoietin (EPO) level, and Hb values <18.5 g/dL in males (range 16.0–18.4) and <16.5 g/dL in females (range 15.0–16.4). Unless bone marrow (BM) histology is carefully integrated in the diagnostic process, a diagnosis of essential thrombocytopenia (ET)\textsuperscript{10,11} or myeloproliferative neoplasm (MPN) unclassifiable (MPN-U) may be made in these cases following the 2008 WHO thresholds. Patients presenting with the lower threshold values and a BM biopsy diagnostic of PV had been designated as so-called masked/prodromal polycythemia vera (mPV)\textsuperscript{12} It has been shown that the BCSH criteria\textsuperscript{6,7} based on Hct appeared as more sensitive than the WHO approach based on Hb values. In fact, using the BCSH criteria\textsuperscript{6,7} the fraction of mPV patients decreased from 35% to 15\%\textsuperscript{13,14} Following proposals of experts,\textsuperscript{15} a Hb level of 16.5 g/dL in men and 16.0 g/dL for women or a Hct level of 49% in men and 48% in women was determined to be the optimal cutoff levels for distinguishing JAK2-mutated ET from mPV\textsuperscript{13} and recently adopted by the 2016 revised WHO classification.\textsuperscript{3} In this regard, correct diagnosis of mPV constitutes an important issue and will certainly improve clinical management of patients while avoiding undertreatment and prevention of thrombotic events.\textsuperscript{9,15}

Given the importance which is now given to an accurate histologic diagnosis of PV for purpose of identifying mPV,\textsuperscript{23} we conducted an international consensus study aimed to examine the interobserver diagnostic reproducibility. To this aim, we assessed the concordance achieved among a panel of six hematopathologists in identifying cases of mPV by BM morphology by separating them from cases of overt PV, ET, and controls.

### 2 | MATERIAL AND METHODS

#### 2.1 | Study design and performance

The current work is the result of a collaborative project by the European LeukemiaNet (ELN) with international experts’ participation. This manuscript was developed through extensive discussions that took place during two meetings in Frankfurt and Amsterdam, respectively. These meetings were led by an expert panel of hematopathologists and clinicians that included authors of the 2008/2016 WHO guidelines, members of the MPN subcommittee of the ELN, as well as recognized international authorities. Pre- and post-meeting inputs from study participants were sought through electronic communications and adjudicated through consensus discussions.

A total of 98 JAK2-positive cases were selected for the blinded review from two different sources: (1) the Frankfurt MPN database from which 57 samples with the final diagnosis of PV as well as control cases (n = 19) were randomly recruited by one of the panelists (H.M.K.). The latter included ET (n = 10), primary myelofibrosis (PMF, n = 5), MPN-U and reactive BM changes (n = 4), and one case with myelodysplastic syndrome (MDS). (2) The New York clinical database, from which 24 samples were selected by another panelist (A.O.), included 22 PV patients and one sample with the diagnosis of ET and one with PMF, respectively. In this cohort Cr-51 red blood cell mass (RCM) measurement were available, as reported in more detail elsewhere.\textsuperscript{3}

The representative, buffered formalin-fixed and paraffin-embedded BM trephine biopsies were cut and sections stained with hematoxylin/eosin (H&E) and silver impregnation for reticulin fibers following Gordon-Sweets or Gomori methods and occasionally Prussian blue staining was applied. Using a multiheaded microscope the six members of the histopathology panel (H.M.K.A.O., J.T., C.B-R., U.G., R.P.H.) evaluated all BM specimens in an independent and blinded fashion (only gender and age were known) and entered their results into a corresponding form sheet during a microscopic review session. The panel applied a combination of histologic features (summarized in Table 1) to assign cases as PV or ET, or others (PMF, MDS, reactive lesions). Degree of inter-rater reliability (IRR) was calculated as overall agreement among panel members based on the number of correct classifications. The 95% confidence intervals for proportions of consensus were computed using normal approximations. Fleiss kappa statistics were used to determine the inter-observer agreement on the final grading. Additionally, we calculated an Intraclass Correlation Coefficient (ICC) that incorporates the magnitude of disagreement between raters, with IRR being excellent for values between 0.75 and 1.0. All computations were performed in R version 3.0.1, using the package “irr” for the Fleiss kappa and ICC calculations.\textsuperscript{16–18}

### 3 | RESULTS

Evaluation of BM morphology among the panelists showed that all 79 PV cases displayed age-adjusted increased hematopoietic cellularity due to increased trilineage proliferation (so-called panmyelosis) associated with the presence of a conspicuous variability in the size of megakaryocytes (high degree of cytologic pleomorphism) without significant nuclear atypia, consistent with PV (Figure 1A,B). After incorporating the relevant Hb threshold values according to the WHO classification\textsuperscript{5} and the originally proposed guidelines for mPV\textsuperscript{12} (males Hb > 16.0; females Hb > 15.0 g/dL), 48 cases were classified as mPV, including the group of 15 patients from the New York cohort shown to have increased RCM. The remaining 31 patients had WHO-defined overt PV (Hb > 18.5 g/dL in men and > 16.5 g/dL in women). The clinical data at presentation for mPV and overt PV are shown in Table 2. A few specimens, particularly of mPV displayed a minor degree of reticulin fibrosis (Figure 1C), and a small amount of stainable iron deposits were detectable in few samples. The other 19 samples included cases of ET and reactive marrow which were also classified correctly. Discrimination of PV cases from ET (Figure 1D) was performed with complete consensus between the panelists according to the key morphological features (Table 1) as was distinction between PV and reactive BM changes (reactive polycythemia and reactive thrombocytosis), MDS and prefibrotic/early PMF (pre-PMF) which were also included among the controls.

In contrast with the PV specimens where a concordance rate of 95% (75/79) was observed, samples of the control group were more difficult to classify, with an unclear discrimination in two cases: one
case with a differential diagnosis of mPV versus ET; the second case differential was mPV versus early stage PMF presenting with increased erythropoiesis. In all patients, the presence of increased hematopoietic cellularity was regarded as critical for the histological diagnosis of mPV. Concerning overall agreement among the six hematopathologists results of concordance of the blinded microscopic assessment during the review.

TABLE 1  Key points to discriminate ET from PV (modified from references #[5], #[19], #[34])

<table>
<thead>
<tr>
<th>ET</th>
<th>PV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cellularity (age-matched)</strong></td>
<td><strong>Hypercellular</strong> (&lt;20% over expected cellularity for age)</td>
</tr>
<tr>
<td>· Normocellular (&lt; 20% over expected cellularity for age)</td>
<td>· Hypercellular (&lt;20% over expected cellularity for age, usually nearly 100%)</td>
</tr>
<tr>
<td><strong>Increased lineage(s)</strong></td>
<td><strong>Erythro-/megakaryo-/granulopoiesis (panmyelosis)</strong></td>
</tr>
<tr>
<td>· Megakaryocytes only</td>
<td>· Left shift in erythro- and granulopoiesis</td>
</tr>
<tr>
<td>· No left shift in erythro-/granulopoiesis</td>
<td></td>
</tr>
<tr>
<td><strong>Morphological characteristics</strong></td>
<td><strong>Mature megakaryocytes with significant variability in size (pleomorphism)</strong></td>
</tr>
<tr>
<td>· Large/giant, mature megakaryocytes with hyperlobulated nuclei</td>
<td></td>
</tr>
<tr>
<td><strong>Bone marrow stroma</strong></td>
<td><strong>Lymphoid nodules rare or absent.</strong></td>
</tr>
<tr>
<td>· Usually no increase in reticulin fibers (&lt;5%)</td>
<td>· Mild increase in reticulin fibers (&lt;20%)</td>
</tr>
<tr>
<td>· Normal sinuses</td>
<td>· Dilated sinuses, some with intraluminal erythrocytes</td>
</tr>
<tr>
<td>· Lymphoid nodules rare or absent.</td>
<td>· Lymphoid nodules up to 20%</td>
</tr>
<tr>
<td><strong>Molecular features</strong></td>
<td><strong>100% JAK2/Exon 12</strong></td>
</tr>
<tr>
<td>· JAK2:64%</td>
<td></td>
</tr>
<tr>
<td>· CALR:15%</td>
<td></td>
</tr>
<tr>
<td>· MPL:4%</td>
<td></td>
</tr>
<tr>
<td>· Triple-negative:16%</td>
<td></td>
</tr>
</tbody>
</table>
| **a**Data from Rotunno et al. [31].

FIGURE 1  A, Overview of so-called masked/prodromal PV. Note the increased age-adjusted marrow cellularity due a proliferation which includes all three cell lineages (panmyelosis). B, High power reveals conspicuous variability in the megakaryocyte cell size (pleomorphism). C, ET in contrast, is characterized by normal age-adjusted marrow cellularity with proliferation limited to the megakaryocytes. In ET there is much less pleomorphism with predominance of large to giant mature and hyperlobulated megakaryocytes. D, In about 10% of mPV such as in this case, a minor increase in reticulin fibers can be observed. A, C: H&E, ×150; B: PAS, ×250; D: Gomori stain, ×150. [Color figure can be viewed at wileyonlinelibrary.com]
session are shown in Table 3. Noteworthy was that more than a 90% agreement could be achieved among the six panelists. Kappa statistics of above 0.8 were in line with the almost perfect agreement among the panelists (Table 3). Altogether, histological BM features of mPV were found to be identical to overt PV.

4 | DISCUSSION

To our knowledge, this is the first consensus conference in which an international panel of hematopathologists has systematically analyzed and validated the key morphologic features associated with a diagnosis of PV. The study results confirm and extend the notion that PV is characterized by specific morphological BM findings by applying the proposed lower threshold values (>16.5 g/dL for men >16.0 g/dL for women). Furthermore, our results demonstrate that the diagnostic morphological criteria defined by the 2016 WHO guidelines are reproducible with a high rate of concordance. Finally, the panel found no significant morphological differences between overt PV and early phase disease including the so-called mPV.

Patients with JAK2-mutated ET share many molecular genetic features with PV and it has been postulated that both groups may belong to a biological continuum in which the extent of erythrocytosis is determined by physiological and genetic factors. In some published series JAK2-mutated ET patients were reported to undergo transformation to PV with a cumulative risk of 29% at 15 years, while no transformation to PV was observed in CALR-mutated ET patients. It was therefore suggested by these authors that JAK2-mutated ET and PV may represent different phenotypes in the evolution of a single MPN. However, the total of 745 ET patients included in the mentioned study had been diagnosed between 1980 and 2012, and all diagnoses were assigned in accordance with criteria in use at the time of initial disease presentation. Since the WHO diagnostic guidelines had only been applied after 2002, a considerable fraction of the total cohort of ET patients would have been diagnosed according to the criteria of the Polycythemia Vera Study Group (PVSG) which are inadequate to capture differences between early PV and ET. Moreover, 30% of patients (142/466) of this cohort displayed evidence of erythrocytosis or leukocytosis about 22 months before the clinical diagnosis of PV, and in these patients the Hb values, white blood cell counts, serum EPO levels, and degree of splenomegaly were above the usual range for ET according to the WHO definition. The exceptionally high rate of transformation into PV in this trial is unusual, because contrasting results were obtained in a strictly WHO-defined cohort of 299 ET patients after almost 13 years of follow-up. In this series the frequency of conversion was reported as 5% in the 159 JAK2-mutated and 3% in the non-mutated cases. Moreover, an even lower incidence was found by Rutunno and coworkers with a rate of only 1.4% in 369 JAK2-mutated ET patients after about 6 years. In aggregate, both studies support the WHO concept of distinctive MPN-subtypes and not the model of a disease continuum in JAK2-mutated MPN. A most likely explanation for this discrepancy is that “pre-PV ET” cases include a high proportion of mPV, particularly among cases recruited before 2002. It is well established that PV, although displaying characteristic BM features, may clinically mimic ET at onset prior to manifesting as overt PV at a later stage. These differences cannot be captured by the PVSG criteria and may be avoided by applying strictly a WHO-defined clinicopathological approach.

Further to strengthen the separation of PV from ET, the diagnostic guidelines of the WHO have been recently revised and modified regarding the major diagnostic criteria. According to this new guideline, a Hb value of >16.5 g/dL in men and >16.0 g/dL for women or a Hct value of >49% in men and >48% in women has ever been reached.

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Hematological data in patients with so-called masked/prodromal PV (mPV), versus overt PV (mean ± 95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mPV</td>
<td>PV</td>
</tr>
<tr>
<td>No. of patients</td>
<td>48</td>
</tr>
<tr>
<td>Hemoglobin (g/dL)</td>
<td>16.2[16.0–16.6]</td>
</tr>
<tr>
<td>Hematocrit (%)</td>
<td>49.6[48.0–51.6]</td>
</tr>
<tr>
<td>White blood cells (×10³/μL)</td>
<td>12.1[10.7–13.5]</td>
</tr>
<tr>
<td>Platelets (×10³/μL)</td>
<td>687[568–806]</td>
</tr>
<tr>
<td>Spleen size (cm³)</td>
<td>0.9[0.3–1.5]</td>
</tr>
<tr>
<td>EPO-subnormal (%)</td>
<td>84</td>
</tr>
<tr>
<td>LDH (U/L)</td>
<td>268[237–298]</td>
</tr>
</tbody>
</table>

Abbreviations: EPO = Erythropoetin; LDH = Lactate dehydrogenase. cm below left costal margin.

of the chosen approach, have shown that a meaningful discrimination between PV and ET may not be achieved without an adequate BM evaluation, an important component of diagnostic algorithms for PV and ET diagnoses in the 2016 revised WHO classification.

Patients with JAK2-mutated ET share many molecular genetic features with PV and it has been postulated that both groups may belong to a biological continuum in which the extent of erythrocytosis is determined by physiological and genetic factors. In some published series JAK2-mutated ET patients were reported to undergo transformation to PV with a cumulative risk of 29% at 15 years, while no transformation to PV was observed in CALR-mutated ET patients. It was therefore suggested by these authors that JAK2-mutated ET and PV may represent different phenotypes in the evolution of a single MPN. However, the total of 745 ET patients included in the mentioned study had been diagnosed between 1980 and 2012, and all diagnoses were assigned in accordance with criteria in use at the time of initial disease presentation. Since the WHO diagnostic guidelines had only been applied after 2002, a considerable fraction of the total cohort of ET patients would have been diagnosed according to the criteria of the Polycythemia Vera Study Group (PVSG) which are inadequate to capture differences between early PV and ET. Moreover, 30% of patients (142/466) of this cohort displayed evidence of erythrocytosis or leukocytosis about 22 months before the clinical diagnosis of PV, and in these patients the Hb values, white blood cell counts, serum EPO levels, and degree of splenomegaly were above the usual range for ET according to the WHO definition. The exceptionally high rate of transformation into PV in this trial is unusual, because contrasting results were obtained in a strictly WHO-defined cohort of 299 ET patients after almost 13 years of follow-up. In this series the frequency of conversion was reported as 5% in the 159 JAK2-mutated and 3% in the non-mutated cases. Moreover, an even lower incidence was found by Rutunno and coworkers with a rate of only 1.4% in 369 JAK2-mutated ET patients after about 6 years. In aggregate, both studies support the WHO concept of distinctive MPN-subtypes and not the model of a disease continuum in JAK2-mutated MPN. A most likely explanation for this discrepancy is that “pre-PV ET” cases include a high proportion of mPV, particularly among cases recruited before 2002. It is well established that PV, although displaying characteristic BM features, may clinically mimic ET at onset prior to manifesting as overt PV at a later stage. These differences cannot be captured by the PVSG criteria and may be avoided by applying strictly a WHO-defined clinicopathological approach. Further to strengthen the separation of PV from ET, the diagnostic guidelines of the WHO have been recently revised and modified regarding the major diagnostic criteria. According to this new guideline, a Hb value of >16.5 g/dL in men and >16.0 g/dL for women or a Hct value of >49% in men and >48% in women has ever been reached.

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Concordance among six panelists concerning PV diagnosis and controls (n = 98)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall agreement</td>
<td>92.6%</td>
</tr>
<tr>
<td>Fleiss Kappa</td>
<td>0.812</td>
</tr>
<tr>
<td>Intraclass correlation (ICC)</td>
<td>0.982</td>
</tr>
</tbody>
</table>
women is sufficient to establish the presence of erythrocytosis while a corresponding BM evaluation is necessary to establish the diagnosis of PV.5

In conclusion, a blinded review by a central panel of six hematopathologists has validated the 2016 revised WHO classification5 regarding the specific BM histological features which characterize PV and confirmed and extended the notion that there are no significant differences regarding BM morphology between mPV and overt stages of PV. Moreover, there was also a high rate of reproducibility regarding the discrimination between PV and ET as well as in relation to diagnosing pre-PMF and reactive conditions.

ACKNOWLEDGMENT
Supported by the Alpine Oncology Foundation (AOF).

CONFLICT OF INTEREST
Nothing to report.

AUTHOR CONTRIBUTIONS
Conception and design: Hans Michael Kvasnicka, Attilio Orazi, Juergen Thiele, Giovanni Barosi, Tiziano Barbiu
Provision of study materials or patients: Hans Michael Kvasnicka, Attilio Orazi, Richard Silver
Collection and assembly of data: Hans Michael Kvasnicka, Attilio Orazi, Richard Silver
Manuscript writing: Hans Michael Kvasnicka, Juergen Thiele, Attilio Orazi,

REFERENCES


