Response criteria for essential thrombocytopenia and polycythemia vera: result of a European LeukemiaNet consensus conference

Giovanni Barosi,1 Gunnar Birgegard,2 Guido Finazzi,3 Martin Griesshammer,4 Claire Harrison,5 Hans Carl Hasselbalch,6 Jean-Jacques Kiladjian,7 Eva Lengfelder,8 Mary Frances McMullin,9 Francesco Passamonti,10 John T. Reilly,11 Alessandro M. Vannucchi,12 and Tiziano Barbui13

1Unit of Clinical Epidemiology and Center for the Study of Myelofibrosis, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico Policlinico S. Matteo, Pavia, Italy; 2Department of Haematology, University Hospital, Uppsala, Sweden; 3Division of Transfusion Medicine, Ospedali Riuniti di Bergamo, Bergamo, Italy; 4Department of Hematology and Oncology, Johannes Wesling Medical Centre Minden, Academic Hospital of the University of Hannover, Minden, Germany; 5Department of Haematology, Guy’s and St Thomas’ National Health Service Foundation Trust, London, United Kingdom; 6Department of Hematology, L. Herlev Hospital, University of Copenhagen, Copenhagen, Denmark; 7Assistance Publique-Hopitaux de Paris, Hospital Avicenne, Service d’Hematologie Clinique, Université Paris 13, Bobigny, France; 8Department of Medicine III, Faculty for Medicine Mannheim, University of Heidelberg, Heidelberg, Germany; 9Centre for Cancer Research and Cell Biology, Queen’s University, Belfast, United Kingdom; 10Division of Hematology, University of Pavia, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico Policlinico San Matteo, Pavia, Italy; 11Molecular Haematology Unit, Division of Molecular and Genetic Medicine, Royal Hallamshire Hospital, Sheffield, United Kingdom; 12Unità Funzionale di Ematologia, Dipartimento di Area Critica Medico Chirurgica, Università degli Studi, Firenze, Italy; and 13Division of Hematology, Ospedali Riuniti di Bergamo, Bergamo, Italy

European experts were convened to develop a definition of response to treatment in polycythemia vera (PV) and essential thrombocythemia (ET). Clinico-hematologic (CH), molecular, and histologic response categories were selected. In ET, CH complete response (CR) was: platelet count less than or equal to 400 × 10^9/L, no disease-related symptoms, normal spleen size, and white blood cell count less than or equal to 10 × 10^9/L. Platelet count less than or equal to 600 × 10^9/L or a decrease greater than 50% was partial response (PR). In PV, CH-CR was: hematocrit less than 45% without phlebotomy, platelet count less than or equal to 400 × 10^9/L, white blood cell count less than or equal to 10 × 10^9/L, and no disease-related symptoms. A hematocrit less than 45% without phlebotomy or response in 3 or more of the other criteria was defined as PR. In both ET and in PV, molecular CR was a reduction of any molecular abnormality to undetectable levels. Molecular PR was defined as a reduction more than or equal to 50% in patients with less than 50% mutant allele burden, or a reduction more than or equal to 25% in patients with more than 50% mutant allele burden. Bone marrow histologic response in ET was judged on megakaryocyte hyperplasia while on cellularity and reticulin fibrosis in PV. The combined use of these response definitions should help standardize the design and reporting of clinical studies. (Blood. 2009;113:4829-4833)

Introduction

Essential thrombocythemia (ET) and polycythemia vera (PV) are disorders of hematopoiesis included within the myeloproliferative neoplasms in the forthcoming revised World Health Organization classification.1 The discovery of acquired recurrent molecular abnormalities in JAK2 (JAK2V617F mutation in exon 14 or mutations, insertions, deletions in exon 12) or MPL (mostly MPLW515L/K) has improved their diagnostic approach.2-4 Virtually all patients with PV have a mutation in JAK2, which is represented by the V617F allele in more than 95% of cases; the frequency of the JAK2V617F mutation is approximately 60% in patients with ET, whereas up to 8% of JAK2V617F unmutated ET patients harbor MPLW515L/K mutation. Several studies have also addressed the relevance of V617F mutational status and of mutated allele burden for clinical presentation and prognosis.5 Finally, these molecular defects hopefully provide a more fruitful therapeutic target. The current therapeutic approach in ET and PV is conservative, aimed at lowering the risk of thrombotic events without exposing the patients to an increased risk of leukemic transformation.6 However, molecular targeted therapies are expected to have the potential to affect the natural course of the disease. Such progress calls for standardization of response criteria to accurately assess the value of new treatment modalities.

The responses of ET and PV to therapies, in terms of disease modification, are evaluated by separately analyzing single clinical, hematologic, or molecular parameters; consequently, different definitions of response have been proposed.7-17 Accordingly, there is a pressing need for the development of a standardized definition for monitoring and assessing treatment responses, based on rigorous, consistent, and feasible criteria, which would be especially suitable for the conduct of clinical research and for comparing the outcome of different clinical trials.

Bearing this in mind, a group of European investigators, sponsored by an European Community Network of Excellence (LeukemiaNet) grant, collaborated to define the quality and degree of response in the 2 disorders. The final goal was to develop a definition of response that would be applicable to future clinical studies. This report represents the recommendations from the LeukemiaNet working group.
These categories should be applied to patients with ET and PV receiving disease-modifying therapies. The 3 categories of response should be used in a cumulative, sequential manner, starting from the clinicohematologic response (which represents the minimal set of criteria to be evaluated for assessing response). The decision to use a composite definition of response (clinicohematologic and molecular response, or clinicohematologic, molecular, and histologic response) will depend on the goal and expectations of therapy.

Methods

The LeukemiaNet project’s response criteria were developed using a multistep process, based on a modified National Institutes of Health approach. A 13-member Expert Panel convened in December 2007 was composed of persons with clinical and/or pathologic experience of myeloproliferative neoplasm and was chaired by a clinician with expertise in clinical epidemiology (G. Barosi). The Expert Panel defined the aims of the project, framed the operative context, and participated in the consensus development process aimed at defining the response definitions.

During the initial meeting, the Expert Panel agreed on the aim of the project: “to develop a definition of response to treatment aimed at modifying the natural history of the disease in ET and PV patients, that considers 3 different categories of response, that is, the clinicohematologic, molecular, and histologic response” (Table 1). Thus, the Expert Panel agreed that a unified definition of response is currently not feasible and that the 3 categories should be used in a cumulative, sequential manner, depending on the goal of the therapy.

The first step of the project was aimed at selecting the criteria in their conceptual terms, worded without any numerical or quantitative attributes. For selecting conceptual criteria, a questionnaire was mailed to each member of the Expert Panel asking them to propose candidate conceptual criteria for the 3 response categories. These criteria were further refined in a Delphi process with a second questionnaire that aimed to achieve a rank among the top choices of candidate criteria according to the principle of clinical relevance, that is, the extent by which any criterion would facilitate the assessment of response to disease-modifying agents. All the questionnaires were returned, the candidate conceptual criteria were ranked according to their priority votes, and the criteria that ranked highest and that received at least 80% consensus were included in the list, forming the core set of conceptual criteria.

The conceptual criteria of the core set were then translated, when appropriate, into their operational terms, that is, worded with quantitative or numerical attributes. For this aim, a third questionnaire was mailed to each member of the Expert Panel asking them to propose candidate operational criteria for each conceptual criterion. All the questionnaires were returned, and a core set of operational criteria was formed.

The next step was to grade the level of response for any of the categories of response and provide a definition of any grade of response using the core set of operational criteria. This part of the process was exploited in a consensus meeting using the nominal group technique. Each member of the Expert Panel was asked to choose, for any particular level of response, which elements of the core set criteria he/she considered should be fulfilled. During this process, the expert’s judgments regarding clinical relevance of the criteria were solicited to facilitate an open discussion and ultimately consensus.

Results

The Expert Panel listed 9 criteria to be included as candidate conceptual criteria for the clinicohematologic response assessment in patients with ET, and 10 criteria in patients with PV. Four clinicohematologic criteria in ET, and 5 in PV, had the highest preference rate from the questionnaire as the core set of conceptual criteria to use during therapy evaluation (Table 2). It was decided not to include complications of the disease, such as thromboembolic events or transformation. Such events would rather be considered endpoints of specific studies. Rather, the criteria were chosen to specifically reflect the proliferative status of the disease. Only one conceptual criterion was selected for defining both molecular and histologic response in ET and in PV (Table 2).

The Expert Panel rephrased the conceptual criteria into operational criteria in the last questionnaire and graded some of them by proposing a definition of complete and partial response of the criteria. During the consensus conference, the Expert Panel used the operational criteria to build the final definitions of complete, partial, and no response to be valid for clinicohematologic and molecular categories, and the definition of response for histologic category.

For the clinicohematologic response definition in ET (Table 3), response of platelet count was of prime relevance and the reduction of platelet count was graded differently in the complete response definition (platelet count $\leq 400 \times 10^9/L$) and in partial response (platelet count $\leq 600 \times 10^9/L$). For the clinicohematologic response definition in PV, hematocrit had prime relevance and was a necessary criterion among others in the definition of complete response and sufficient alone for a partial response (Table 3). However, the panel realized that the current hematocrit target of 45% is largely based on prevailing standard of practice that is not evidence-based and that may be modified by future ongoing studies.

| Table 1. Categories of response in ET and PV patients |
|-------------------------------|------------------|
| **Response category**         | **Clinicoheamatologic**                                   |
| Clinicohematologic            | Reduction/normalization of platelet count                 |
| Red blood cell count          | Reduction/normalization of red blood cell count           |
| White blood cell count        | Reduction/normalization of white blood cell count         |
| Bone marrow histologic        | Resolution of disease-related symptoms*                  |
| Bone marrow histologic        | Resolution of disease-related symptoms*                  |
| Histologic                    | Reduction of any specific molecular abnormalities         |
| Bone marrow histologic        | Bone marrow histologic remission defined as the disappearance of megakaryocyte hyperplasia |
| | Bone marrow histologic remission defined as the appearance of age-adjusted normocellularity and no reticulin fibrosis |

*Disease-related symptoms: microvascular disturbances, pruritus, headache.
The definition of molecular response was based on quantitative allele burden of the specific molecular abnormalities (JAK2 V617F, exon 12 mutations, MPL mutations) and took into account that sensitivity of detection varies according to the method used and that a large range of variation in consecutive samples from an individual patient is possible even without therapy. As a consequence, complete molecular response was defined on the basis of the detection levels and partial response was applied only to patients with a baseline value of mutant allele burden greater than 10% (Table 4).

Definitions of histologic response were based on megakaryocyte hyperplasia for ET and cellularity and reticulin fibrosis for PV (Table 5). The reference values for such a definition were taken from the EUMNET response criteria in myelofibrosis. Those criteria agreed that the basic requirement for assessment of cellularity is a representative biopsy, defined as an artifact-free, nontangential sample at least 1.5 cm in length. In addition, the optimal thickness of the paraffin sections should be between 3 and 4 μm, and the cellularity should be documented in relation to age and with respect to normal ranges. Quantity of the fiber content should be determined only in areas of hematomatosis using a scoring system comprising 4 grades.

### Discussion

In this work, we provide response definitions that are valuable for assessing the clinical outcomes of different therapeutic strategies aimed at modifying the disease course in ET and PV. In the absence of a unique and specific biologic marker, a definition of response in ET and PV remains a complex issue, necessitating the incorporation of multiple criteria. The task is further complicated by the paucity of studies reporting scientific evidence on test performance and sensitivity to change of the criteria that could be used for measuring the response. Thus, searching for a definition of response to therapies raised a complex decision, with the pending drawbacks of subjectivity and the arbitrary nature of the resulting criteria. To focus on the problem in a robust manner, we used group techniques and consensus methodologies. The theoretical value of the experts’ consensus approach to influencing definitions is the assumption that such acknowledged experts have an implicit and comprehensive mastery of scientific and practical information that would yield the most appropriate decision. A multistep approach was adopted to help reduce a complex problem into small, easily managed parts, ensuring that all important considerations are being processed and integrating multiple viewpoints into the decision-making process in an explicit and unbiased manner.

Ideally, response criteria could be comprehensive and able to cover all treatment modalities, using the same criteria for all studies. However, there is a need for response criteria to different types of treatments with different goals. Therefore, the Expert Panel found that one unique definition integrating clinicohematologic, molecular, and histologic responses would not fill the need for all studies. The Expert Panel preferred to maintain the possibility of modulating the extent by which a response could be evaluated on the goal of therapy, the mechanism of the drug, and the expected efficacy of the strategy. The present recommendations were proposed also with the intention of preserving some similarities to the response categories in myelofibrosis.

The definition of clinicohematologic response in both ET and PV was constructed on criteria widely used in different definitions of response in these disorders previously reported in the literature but not in this precise combination. The results of the consensus conference suggest that ET patients, when evaluated for the clinicohematologic response, should be assessed according to the variations of 4 major criteria: platelet count, disease-related symptoms, spleen size, and white blood cell count. In PV, 5 major criteria were generated: hematocrit, platelet count, white blood cell count, spleen size, and disease-related symptoms. In both diseases, criteria were selected that were considered markers of proliferative activity of the disease. Moreover, platelet and hematocrit value threshold for defining complete responses in ET and PV, respectively, were selected according to the targets used in the current guidelines for treatment. This implies uncertainty that was

### Table 3. Definition of clinicohematologic response in ET and PV

<table>
<thead>
<tr>
<th>Response grade</th>
<th>Essential thrombocythemia</th>
<th>Polycythemia vera</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete response</td>
<td>(1) Platelet count ≤ 400 × 10^9/L, AND</td>
<td>(1) Hematocrit &lt; 45% without phlebotomy AND</td>
</tr>
<tr>
<td></td>
<td>(2) no disease-related symptoms, AND</td>
<td>(2) platelet count ≤ 400 × 10^9/L AND</td>
</tr>
<tr>
<td></td>
<td>(3) normal spleen size on imaging, AND</td>
<td>(3) white blood cell count ≤ 10 × 10^9/L AND</td>
</tr>
<tr>
<td></td>
<td>(4) white blood cell count ≤ 10 × 10^9/L</td>
<td>(4) normal spleen size on imaging AND</td>
</tr>
<tr>
<td>Partial response</td>
<td>In patients who do not fulfill the criteria for complete response, platelet count ≤ 600 × 10^9/L OR decrease &gt; 50% from baseline</td>
<td>In patients who do not fulfill the criteria for complete response, hematocrit &lt; 45% without phlebotomy OR response in 3 or more of the other criteria</td>
</tr>
<tr>
<td>No response</td>
<td>Any response that does not satisfy partial response</td>
<td>Any response that does not satisfy partial response</td>
</tr>
</tbody>
</table>

* Applies only to patients with a baseline value of mutant allele burden greater than 10%.

### Table 4. Definition of molecular response in ET and PV

<table>
<thead>
<tr>
<th>Response grade</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete response</td>
<td>Reduction of any specific molecular abnormality to undetectable levels</td>
</tr>
<tr>
<td>Partial response*</td>
<td>(1) A reduction of ≥ 50% from baseline value in patients with &lt; 50% mutant allele burden at baseline OR (2) reduction of ≥ 25% from baseline value in patients with &gt; 50% mutant allele burden at baseline.</td>
</tr>
<tr>
<td>No response</td>
<td>Any response that does not satisfy partial response</td>
</tr>
</tbody>
</table>

* Applies only to patients with a baseline value of mutant allele burden greater than 10%.

### Table 5. Criteria for histologic response in ET and PV patients

<table>
<thead>
<tr>
<th>Essential thrombocythemia</th>
<th>Polycythemia vera</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bone marrow histologic remission</td>
<td>Bone marrow histologic remission</td>
</tr>
<tr>
<td>Absence of megakaryocyte hyperplasia</td>
<td>Presence of age-adjusted normocellularity and no reticulin fibrosis</td>
</tr>
</tbody>
</table>
extensively dissected by the Expert Panel. Neither a target of 400 × 10^9/L for platelet count nor a hematocrit of 45% is an unquestionable target for therapy. A value of hematocrit falling in the normal range as a therapeutic target derives from widely cited papers on a relation between the increased hematocrit values and the risk of thrombosis,26,27 but these associations were not subjected to multivariate analysis. Recently, the European Collaboration on Low-Dose Aspirin in Polycythemia Vera study,28 a randomized controlled study comparing aspirin versus placebo in PV, documented that only 48% of patients had hematocrit values below the recommended threshold, whereas 39% and 13% of patients remained between 45% and 50%, and greater than 50%, respectively. Moreover, a multivariate analysis considering all the confounders failed to show any correlation between these hematocrit values and thrombosis. The recognition of this uncertainty has prompted the launch of a prospective, randomized clinical study. In this project, the Expert Panel used a criterion of caution and has maintained the attainment of the lowest threshold of both hematocrit and platelet count as response criteria.

Molecular aberrations in ET and PV have a major biologic and prognostic relevance. At least some of the therapeutic agents in use today have proven capable of reducing the allele burden of JAK2 V617F mutation both in ET and PV.17,29 Furthermore, we expect that the new anti-JAK2 agents could result in a marked reduction in the burden of mutated alleles. A concern of the Expert Panel in providing definition of molecular response was the limitations of detection sensitivity of the currently used methods for the allele burden and the sensitivity to change of the parameter during therapies. The Expert Panel provided a definition of molecular response that highlighted the importance of molecular methods used in the evaluation of response to treatment. Indeed, the sensitivity of current methods may vary substantially; and in any case, no method can detect the allele burden at very low levels. For this reason, the use of the term “complete molecular response” should not be interpreted as an equivalent to complete eradication of mutant cellular clone, and the term “undetectable level” better describes the most complete molecular remission recorded by current methodologies.

Monitoring the changes in bone marrow histology after treatment is not routine in clinical practice and in clinical trials in ET and PV. However, the Expert Panel argued that drugs that promise to change the natural course of the disease need to be assessed for their impact on histopathology of the bone marrow. The hyperplasia of megakaryocytes in ET and hypercellularity and reticulin fibrosis in PV were the parameters the Expert Panel decided to monitor for evaluating the response. A recently published description of scoring of bone marrow fiber density and cellularity could be used as a reference measure.21 Regarding hyperplasia of megakaryocytes, a quantitative definition was not provided, and the Expert Panel accepted a qualitative concept that should be individually graded by the pathologist evaluating the biopsies.

We think that the response criteria presented in this paper are a promising new tool for assessing therapeutic outcomes in patients with ET and PV. These criteria will provide a means to compare the results from different patient cohorts and are recommended to facilitate communication within the scientific community.

Acknowledgments

This work was supported by a grant from Merck, Sharp & Dohme (Hoddesdon, United Kingdom) and Incyte Pharmaceuticals (Wilmington, DE). This work was based on an initiative of the Chronic Myeloproliferative Working Party of the European LeukemiaNet.

Authorship

Contribution: G. Barosi designed the research, chaired the consensus meetings, was on the Expert Panel, and wrote the paper; T.B. designed the research, was on the Expert Panel, and reviewed the paper; and G. Birgegard, G.F., M.G., C.H., H.C.H., J.-J.K., E.L., M.F.M., F.P., J.T.R., and A.M.V. were on the Expert Panel and reviewed the paper.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Giovanni Barosi, Laboratorio di Epidemiologia Clinica, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico Policlinico S. Matteo, Viale Golgi 19, 27100 Pavia, Italy; e-mail: barosig@smatteo.pv.it.

References


From www.bloodjournal.org by guest on March 30, 2017. For personal use only.
Response criteria for essential thrombocythemia and polycythemia vera: result of a European LeukemiaNet consensus conference


Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036. Copyright 2011 by The American Society of Hematology; all rights reserved.