

Inter-observer variance with the diagnosis of myelodysplastic syndromes (MDS) following the 2008 WHO classification

P. Font · J. Loscertales · C. Benavente · A. Bermejo ·
M. Callejas · L. Garcia-Alonso · A. Garcia-Marcilla ·
S. Gil · M. Lopez-Rubio · E. Martin · C. Muñoz ·
P. Ricard · C. Soto · P. Balsalobre · A. Villegas

Received: 29 June 2012 / Accepted: 21 August 2012 / Published online: 5 September 2012
© Springer-Verlag 2012

Abstract Morphology is the basis of the diagnosis of myelodysplastic syndromes (MDS). The WHO classification offers prognostic information and helps with the treatment decisions. However, morphological changes are subject to potential inter-observer variance. The aim of our study was to explore the reliability of the 2008 WHO classification of MDS, reviewing 100 samples previously diagnosed with MDS using the 2001 WHO criteria. Specimens were collected from 10 hospitals and were evaluated by 10 morphologists, working in five pairs. Each observer evaluated 20

samples, and each sample was analyzed independently by two morphologists. The second observer was blinded to the clinical and laboratory data, except for the peripheral blood (PB) counts. Nineteen cases were considered as unclassified MDS (MDS-U) by the 2001 WHO classification, but only three remained as MDS-U by the 2008 WHO proposal. Discordance was observed in 26 of the 95 samples considered suitable (27 %). Although there were a high number of observers taking part, the rate of discordance was quite similar among the five pairs. The inter-observer

P. Font (✉) · P. Balsalobre
Department of Hematology,
Hospital General Universitario Gregorio Marañón,
Madrid, Spain
e-mail: pfontlopez@yahoo.es

J. Loscertales
Department of Hematology, Hospital Universitario de La Princesa,
Madrid, Spain

C. Benavente · A. Villegas
Department of Hematology, Hospital Clinico San Carlos,
Madrid, Spain

A. Bermejo
Department of Hematology,
Hospital Universitario de Fuenlabrada,
Madrid, Spain

M. Callejas · M. Lopez-Rubio
Department of Hematology,
Hospital Universitario Principe de Asturias,
Madrid, Spain

L. Garcia-Alonso
Department of Hematology, Hospital Universitario de Getafe,
Madrid, Spain

A. Garcia-Marcilla · E. Martin
Department of Hematology, Hospital Universitario 12 de Octubre,
Madrid, Spain

S. Gil
Department of Hematology,
Hospital Universitario Puerta de Hierro,
Madrid, Spain

C. Muñoz
Department of Hematology, Hospital Universitario Infanta Leonor,
Madrid, Spain

P. Ricard
Department of Hematology,
Hospital Universitario Fundacion Alcorcon,
Madrid, Spain

C. Soto
Department of Hematology, Fundación Jiménez Díaz,
Madrid, Spain

concordance was very good regarding refractory anemia with excess blasts type 1 (RAEB-1) (10 of 12 cases, 84 %), RAEB-2 (nine of 10 cases, 90 %), and also good regarding refractory cytopenia with multilineage dysplasia (37 of 50 cases, 74 %). However, the categories with unilineage dysplasia were not reproducible in most of the cases. The rate of concordance with refractory cytopenia with unilineage dysplasia was 40 % (two of five cases) and 25 % with RA with ring sideroblasts (two of eight). Our results show that the 2008 WHO classification gives a more accurate stratification of MDS but also illustrates the difficulty in diagnosing MDS with unilineage dysplasia.

Keywords Myelodysplastic syndromes (MDS) · Morphology · WHO classification

Introduction

Myelodysplastic syndromes (MDS) are a heterogeneous group of diseases, presenting with cytopenia and dysplastic changes in bone marrow (BM) and peripheral blood (PB), and a variable risk of transformation to acute myeloid leukemia [1]. Because of the wide variation in clinical outcome, the classification of MDS is still a challenge. In 2008, a new edition of the WHO classification of MDS was published, including some changes from the previous edition published in 2001 [2, 3]. As important aspects, the revised classification allows for a more precise definition of patients previously considered “unclassifiable.” The category of refractory anemia (RA) has been expanded to refractory cytopenia with unilineage dysplasia (RCUD), the category of refractory cytopenia with multilineage dysplasia and ringed sideroblasts (RCMD-RS) is incorporated in RCMD, and the prognostic significance of increased blasts in PB is emphasized, with the change of definition of refractory anemia with excess blasts (RAEB-1) [4].

As all new systems, it is essential to demonstrate whether the categories from the 2008 WHO classification are easy to identify and if the decision of including an individual case in each category is based in reproducible aspects.

There are very few publications evaluating discrepancies in the diagnosis of MDS. Howe et al. reported an excellent correlation among three observers studying 103 samples from patients diagnosed with MDS [5]. However, the three pathologists belonged to the same center. Very recently, Naqvi et al. published their results finding a discrepancy in the morphological diagnosis of MDS between the MDACC and the referring centers [6]. The rate of discrepancy, 12 %, demonstrates the complexity of diagnosing MDS.

The present study was designed with two objectives: First, we explored whether the 2008 WHO classification is more efficient than the previous edition of 2001 in diagnosing 100 patients with dysplastic changes. The second objective was to assess if the 2008 WHO classification allows a good concordance among several hematologists, by distinguishing all the current categories of MDS.

Methods

A total of 100 BM samples from patients previously diagnosed with MDS using the 2001 WHO classification were evaluated in a clinical setting. The study was performed in ten hospitals, and each hematologist provided ten cases. Selection of the cases was made following the criteria of each investigator, and the proportion of samples from each category was not predefined.

The morphological review was done by the same ten hematologists, working in five pairs. First, the observers selected the samples from their centers and reviewed these, using both 2001 and 2008 WHO classification. Later, the specimens were sent to a central coordinator, who exchanged the samples among the observers in a reciprocal manner. This means that each observer evaluated 20 samples, and each sample was analyzed independently by two morphologists. The second observer was blinded to the clinical and laboratory data, except for the PB counts (Fig. 1). The study was approved by the central Ethical Committee of Clinical Research in accordance with the Helsinki Declaration.

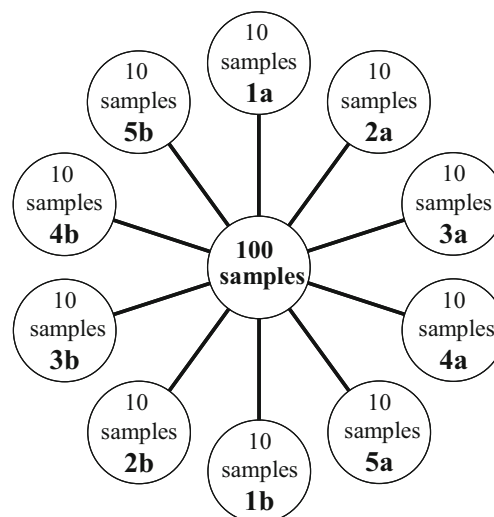


Fig. 1 One hundred cases previously diagnosed with MDS were evaluated by 10 hematologists working in five pairs. Each observer evaluated 20 samples, and each sample was analyzed independently by two morphologists. The specimens were exchanged in reciprocal manner by a central coordinator

Morphological analysis

The MDS samples consisted of BM smears from aspirate, with their correspondent PB smears. Specimens were stained with Wright in 80 cases and May–Grünwald–Giemsa in 20 cases. For evaluation of ring sideroblasts (RS), smears were stained with Prussian blue. The degree of dysplasia in the three hematopoietic lineages and the percentage of blasts and RS were reported according to the criteria of the 2001 and 2008 WHO classification. At least 500 nucleated BM cells were counted.

Cytogenetics

Cytogenetics was available in 65 cases. Cytogenetic aberrations were grouped according to the original IPSS [7]. Cytogenetic data will not be shown as it makes no difference for the purpose of the present study.

Statistical methods

The interobserver concordance was evaluated using the Cohen κ test.

Results

All the samples were diagnosed originally by the FAB [8] and the 2001 WHO criteria. The first step was to reclassify the samples in the referring centers by the 2008 WHO classification. Table 1 shows the cases that were reclassified according with the new definitions by the 2008 WHO classification.

Table 1 Cases of MDS reclassified by the referring observers according the 2008 WHO

| 2008 WHO | 2001 WHO classification | | | | |
|---------------------|-------------------------|----------------------|-------------------------|-----------------------|-----------------------|
| | RA <i>n</i> =1 | RCMD <i>n</i> =31 | RCMD-RS <i>n</i> =12 | RAEB1 <i>n</i> =11 | MDS-U <i>n</i> =19 |
| RCUD, <i>n</i> =6 | 1 | | | | 5 |
| RCMD, <i>n</i> =52 | | 29 | 11 | | 12 |
| RAEB1, <i>n</i> =13 | | 2 | | 11 | |
| MDS-U, <i>n</i> =3 | | | 1 | | 2 |

The table shows the reclassified diagnoses, according to the new definitions of RCUD, RCMD, RAEB-1, and MDS-U by the 2008 WHO

RCUD refractory cytopenia with unilineage dysplasia, *RA* refractory anemia, *RCMD* refractory cytopenia with multilineage dysplasia, *RCMD-RS* refractory cytopenia with multilineage dysplasia and ringed sideroblasts, *RAEB1* RA with excess blasts type 1, *MDS-U* unclassified-MDS

The percentage of myeloblasts in PB led to change the diagnosis of three cases with multilineage dysplasia. Two cases of RCMD were classified as RAEB-1 because they showed 2 % blasts in PB. One case of RCMD-RS was defined as unclassified MDS (MDS-U) by the 2008 WHO classification, with 1 % of myeloblasts in the PB.

Nineteen cases were considered as MDS-U by the 2001 WHO classification. Twelve cases were reclassified using the 2008 WHO classification as RCMD and five as RCUD. Finally, only three cases remained as MDS-U by the 2008 WHO proposal.

After this, the samples were exchanged for the second review. Five samples were rejected by the second observers because of bad quality of the BM or PB smears, and 95 samples were considered suitable for the study. Discordance was observed in 26 samples (27 %). The inter-observer agreement for all the series was 73 % (kappa test 0.59).

Regarding the five pairs of observers, all of them found some discrepancies (Table 2). The percentage of discordance was 21 % by the first pair (four discrepancies of 19 suitable samples), 26 % by the second pair (five of 19), 26 % by the third pair (five of 19), 33 % by the fourth pair (six of 18), and 30 % by the fifth (six of 20). The 26 cases with interobserver discrepancy are shown in Table 3.

Since the results by pairs were quite homogeneous, we considered the series as a whole, to study the rate of agreement by categories. First, we focused in the enumeration of RS and blast cells. There were 20 cases with 15 % or more of RS. All these cases were easily identified by the second observers. Regarding the study of blast cells, we divided the series in two groups: The first group included the cases with increased number of blasts in BM (≥ 5 %) or PB (≥ 2 %); it comprised 22 cases diagnosed as RAEB-1+RAEB type 2 (RAEB-2) by the referring centers. The second group included the cases with MDS and without excess blasts (73 cases). All the referring diagnoses with an excess of blasts were confirmed, and only six cases with less than 5 % of BM blasts were diagnosed as RAEB-1 by the second observers. The inter-observer agreement regarding specimens with and without excess blasts was 94 %.

Then we studied the inter-observer concordance by the 2008 WHO specific categories (Table 4). Concordance was detected in 84 % of cases with RAEB-1 (10 of 12 cases), 90 % with RAEB-2 (nine of 10), 100 % of cases with myelodysplastic/myeloproliferative neoplasms (MDS/MPN) (four of four), and 100 % with MDS-U (three of three).

The rate of concordance of cases with RCMD (74 %) was quite similar among the five pairs of observers

Table 2 Distribution of samples among the five pairs of observers, regarding 2008 WHO classification

| | 2008 WHO classification | | | | | | | |
|-----------------------------|-------------------------|---------------------|----------------------|-----------------------|-----------------------|----------------------|----------------------|------------------------|
| | RCUD <i>n</i> =5 | RARS <i>n</i> =8 | RCMD <i>n</i> =50 | RAEB1 <i>n</i> =12 | RAEB2 <i>n</i> =10 | Del5q <i>n</i> =3 | MDS-U <i>n</i> =3 | MDS/MPN <i>n</i> =4 |
| Pair 1. 19 suitable samples | 3 | 1 | 7 | 3 | 4 | 0 | 1 | 0 |
| Rate of discordance 4/19 | (1) | | (2) | | (1) | | | |
| Pair 2. 19 suitable samples | 1 | 3 | 4 | 4 | 3 | 2 | 0 | 2 |
| Rate of discordance 5/19 | (1) | (2) | | (1) | | (1) | | |
| Pair 3. 19 suitable samples | 0 | 0 | 15 | 1 | 2 | 0 | 1 | 0 |
| Rate of discordance 5/19 | | | (5) | | | | | |
| Pair 4. 18 suitable samples | 0 | 3 | 10 | 1 | 1 | 1 | 0 | 2 |
| Rate of discordance 6/18 | | (3) | (3) | | | | | |
| Pair 5. 20 suitable samples | 1 | 1 | 14 | 3 | 0 | 0 | 1 | 0 |
| Rate of discordance 6/20 | (1) | (1) | (3) | (1) | | | | |

The cases with discrepancies are showed in brackets

(Table 2). They noted discrepancies in 13 of 50 samples: Five samples were identified as RAEB-1 and eight as RCUD. Multilineage dysplasia was recognized in 42 of 50 cases (84 % agreement). The observers only found concordance in two of five cases with RCUD (40 %), in two of eight cases with refractory anemia with ring sideroblasts (RARS) (25 %), being the discordance due to multilineage dysplasia considered by the second observer (Table 3).

Discussion

The fourth edition of the WHO classification includes significant changes in order to improve the previous system published by the WHO in 2001. One important characteristic of a good taxonomy system is reproducibility with little inter-observer variation. This is essential regarding MDS, as the current classification offers

prognostic information that can guide the treatment. The objective of our study was to explore the reliability of the 2008 WHO, compared with the 2001 WHO classification, among several observers from ten institutions. When the observers recoded their own samples according to the 2008 WHO classification, we found that the 2008 WHO was able to diagnose in a specific category 17 of 19 cases that were previously “unclassified” by using the 2001 classification.

Most of the cases coded as “unclassified” by the 2001 WHO showed only one cytopenia with dysplasia in two or three lineages. The 2008 WHO classification identifies these cases as RCMD. These patients with one cytopenia in the context of multilineage dysplasia seem to have a longer survival than patients with two cytopenias [9]. Anyhow, the 2008 WHO classification gives a clearer definition of the RCMD, and it enables a more accurate stratification of patients in our series.

After this reclassification, the specimens were exchanged in a reciprocal manner among the group. Although we had a large number of observers taking part in our study, the concordance among the different pairs was very similar. The higher rate of discrepancy (33 %) was detected by the pair of observers who included more cases without excess blasts.

We noted that, in cases with morphological discrepancy, the second observer usually assessed a higher risk category than the referring. Eight cases with unilineage dysplasia were identified as RCMD. Five cases with RCMD and one with del5q were defined as RAEB-1, and two cases with RAEB-1 were reclassified as RAEB-2. Probably, there was a bias of over-diagnosis because all the participants knew that they were studying samples previously diagnosed with MDS.

Table 3 Twenty-six cases with interobserver discrepancies regarding 2008 WHO classification

| Referring diagnosis (<i>n</i> cases) | Review (<i>n</i> cases) |
|---------------------------------------|------------------------------|
| RCUD (3) (1 RN ^a , 2RA) | RCMD (2) RA (1) |
| RARS (6) | RCMD (6) |
| RCMD (13) | RARS (3) RCUD (5) RAEB-1 (5) |
| RAEB1 (2) | RAEB-2 (2) |
| RAEB2 (1) | RAEB-1(1) |
| MDS del(5q) (1) | RAEB-1 (1) |

^a The case classified as RN was coded as RA by the second observer. This patient showed trisomy 8

RN refractory neutropenia, RA refractory anemia

Table 4 Agreement by second observers regarding all categories of 2008 WHO classification

| 2008 WHO categories | RCUD | RARS | RCMD | RAEB1 | RAEB2 | Del 5q | MDS-U | MDS/MPN |
|--|----------|----------|-----------|-----------|----------|----------|-----------|-----------|
| Diagnosis by referring observers, <i>N</i> =95 | 5 | 8 | 50 | 12 | 10 | 3 | 3 | 4 |
| Agreement by 2nd observers (%) | 2 (40 %) | 2 (25 %) | 37 (74 %) | 10 (84 %) | 9 (90 %) | 2 (66 %) | 3 (100 %) | 4 (100 %) |

Regarding patients with or without excess blasts (threshold ≥ 5 % BM blasts), the inter-observer agreement was excellent (94 %). Other authors also have demonstrated very good concordance in recognizing blasts cells [5, 6]. Detection of BM blasts by morphology is essential for defining prognosis. The threshold of 5 % of BM blasts was defined by the FAB system in 1982, and it is still valid in the context of WHO classification. There was also a good agreement between the categories RAEB-1 and RAEB-2. In 2008, the International Working Group of Morphology of MDS developed standardized definitions for RS, myeloblasts, and promyelocytes [10]. These clear definitions contribute to improve precision in the diagnosis of MDS.

The category of RCMD comprised the highest number of samples because it included most of patients with a previous diagnosis of MDS-U and RCMD-RS. The definition of RCMD requires dysplastic changes in ≥ 10 % cells in two or more myeloid lineages. Multilineage dysplasia, in the absence of BM blasts, is associated with worse prognosis than unilineage dysplasia [11]. Matsuda et al. have identified some morphological features, like pseudo-Pelger–Huet anomalies, or micromegakaryocytes, as unfavorable prognostic factors [12]. But this information is not modifying in the current WHO classification, so we considered an agreement when both observers assessed dysplasia in two or more lineages, regardless of the lineages affected. The study showed 74 % agreement among the second observers in this category. But the interobserver concordance was higher (84 %) when we focused on the detection of multilineage dysplasia, independently of the enumeration of blasts. Dysplastic changes in two or more lineages were easily recognized in 42 cases. This consistent agreement was expected, as RCMD is a good example of clonal hematopoietic stem cell disease, with evident dysplasia in at least two BM lines.

Few cases showed unilineage dysplasia (RCUD, RARS) in our series. It seems reasonable, as these categories are less frequent than RCMD or RAEB, and we did not define a number of cases by categories.

However, the low agreement among observers is striking. Probably, these categories are misdiagnosed. The arbitrary threshold of 10 % of dysplasia in each cell line might turn out quite low, when we are reviewing samples that we know

previously diagnosed with MDS. Other authors reported similar results. Howe et al. showed a very good agreement in their interobserver variation study, but they found discrepancy in four of nine cases with RA according 2001 WHO classification. Furthermore, there were only eight cases with discrepancy in their series: Seven showed unilineage dysplasia (four RA, three RARS) [5].

Within patients with unilineage dysplasia, cases diagnosed with RARS by the referring observers also showed high discrepancy. The second observer found dysplasia ≥ 10 % in neutrophil precursors or megakaryocytes in six out of eight cases. The distinction between unilineage and multilineage dysplasia in the RA and RARS categories is very important, as Germing et al. demonstrated in their retrospective and prospective series [13, 14]. The negative impact of multilineage dysplasia, confirmed by these authors, led to the merging of the entities RCMD and RCMD with ringed sideroblasts in the 2008 WHO classification. In concordance with Germing and colleagues, Patnaik et al. also showed that the percentage of ring sideroblasts per se does not influence the prognosis of MDS without excess of blasts, but the diagnosis of RARS was one of the independent indicators of overall survival [15].

Unilineage dysplasia by the WHO has been associated with a favorable prognosis, and it is one of the factors that influence the WPSS [11]. This system allocates a different risk, based on the subjective criteria of the observer who makes the diagnosis. The IPSS and the risk model developed by Kantarjian et al. are not influenced by the degree of dysplasia, so all of these cases with discrepancies showed the same risk by these scores [16]. Although our series had a few number of these cases, we believe that our study illustrates the difficulty of diagnosing unilineage dysplasia.

We conclude that our study shows a good agreement in detecting cases with or without excess blasts and also in detecting multilineage dysplasia that are very relevant prognostic factors. However, the diagnosis of unilineage dysplasia was not reproducible. A larger series of patients is needed to confirm these results, but we stress the importance of carefully making the diagnosis of unilineage MDS. In the future, new prognostic tools, as molecular and immunophenotypic findings, would help to understand the current morphological discrepancies.

Acknowledgments Celgene provided a grant to the Asociación Madrileña de Hematología y Hemoterapia (AMHH) for the courier service. The authors would like to thank Angel Cedillo (AMHH) for his help in the exchange of samples among the hospitals and Jose María Bellón for the statistical analysis. We thank Lawrence JC Baron for providing language support and Dolores Subirá for her useful comments.

Conflict of interest The authors declare that they have no conflict of interest.

References

- Tefferi A, Vardiman JW (2009) Myelodysplastic syndromes. *N Engl J Med* 361:1872–1885
- Brunning RD, Orazi A, Germing U, Le Beau MM, Porwit A, Baumann I, Vardiman JW, Hellstrom-Lindberg E (2008) Myelodysplastic syndromes/neoplasms, overview. In: Swerdlow SH, Campo E, Harris NL et al (eds) World Health Organization classification of tumours of haematopoietic and lymphoid tissues, 4th edn. IARC, Lyon, pp 88–93
- Brunning RD, Bennett JM, Flandrin G, Matutes E, Head D, Vardiman JW, Harris NL (2001) Myelodysplastic syndromes. Introduction. In: Jaffe ES, Harris NL, Stein H, Vardiman JW (eds) World Health Organization classification of tumours. Pathology and genetics of tumours of haematopoietic and lymphoid tissues. IARC, Lyon, pp 63–67
- Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A et al (2009) The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood* 114:937–951
- Howe RB, Porwit-MacDonald A, Wanat R, Tehranchi R, Hellström-Lindberg E (2004) The WHO classification of MDS does make a difference. *Blood* 103:3265–3270
- Naqvi K, Jabbour E, Bueso-Ramos C, Pierce S, Borthakur G, Estrov Z et al (2011) Implications of discrepancy in morphologic diagnosis of myelodysplastic syndrome between referral and tertiary care centers. *Blood* 118:4690–4693
- Greenberg P, Cox C, LeBeau MM, Fenaux P, Morel P, Sanz G et al (1997) International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* 89:2079–2088
- Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, Sultan C (1982) French–American–British (FAB) Cooperative Group. Proposals for the classification of the myelodysplastic syndromes. *Br J Haematol* 51:189–199
- Verburgh E, Achten R, Louw VJ, Brusselmans C, Delforge M, Boogaerts M et al (2007) A new disease categorization of low-grade myelodysplastic syndromes based on the expression of cytopenia and dysplasia in one versus more than one lineage improves on the WHO classification. *Leukemia* 21:668–677
- Mufti GJ, Bennett JM, Goasguen J, Bain BJ, Baumann I, Brunning R et al (2008) Diagnosis and classification of myelodysplastic syndrome: International Working Group on Morphology of Myelodysplastic Syndrome (IWGM-MDS) consensus proposals for the definition and enumeration of myeloblasts and ring sideroblasts. *Haematologica* 93:1712–1717
- Malcovati L, Germing U, Kuendgen A, Della Porta MG, Pascutto C, Invernizzi R et al (2007) Time-dependent prognostic scoring system for predicting survival and leukemic evolution in myelodysplastic syndromes. *J Clin Oncol* 25:3503–3510
- Matsuda A, Germing U, Jinnai I, Iwanaga M, Misumi M, Kuendgen A et al (2007) Improvement of criteria for refractory cytopenia with multilineage dysplasia according to the WHO classification based on prognostic significance of morphological features in patients with refractory anemia according to the FAB classification. *Leukemia* 21:678–686
- Germing U, Gattermann N, Strupp C, Aivado M, Aul C (2000) Validation of the WHO proposals for a new classification of primary myelodysplastic syndromes: a retrospective analysis of 1600 patients. *Leuk Res* 24:983–992
- Germing U, Strupp C, Kuendgen A, Isa S, Knipp S, Hildebrandt B et al (2006) Prospective validation of the WHO proposals for the classification of myelodysplastic syndromes. *Haematologica* 91:1596–1604
- Patnaik MM, Hanson CA, Sulai NH, Hodnefield JM, Knudson RA, Ketterling RP et al (2012) Prognostic irrelevance of ring sideroblast percentage in World Health Organization defined myelodysplastic syndromes without excess blasts. *Blood* 119:5674–5677
- Kantarjian H, O'Brien S, Ravandi F, Cortes J, Shan J, Bennett JM et al (2008) Proposal for a new risk model in myelodysplastic syndrome that accounts for events not considered in the original International Prognostic Scoring System. *Cancer* 113:1351–1361