

Artificial intelligence to empower diagnosis of myelodysplastic syndromes by multiparametric flow cytometry

Valentin Clichet,¹ Delphine Lebon,^{2,3} Nicolas Chapuis,⁴ Jaja Zhu,⁵ Valérie Bardet,⁵ Jean-Pierre Marolleau,^{2,3} Loïc Garçon,^{1,3} Alexis Caulier^{2,3,6,7} and Thomas Boyer^{1,3}

¹Service d'Hématologie Biologique, CHU Amiens-Picardie, Amiens, France; ²Service d'Hématologie Clinique et de Thérapie Cellulaire, CHU Amiens-Picardie, Amiens, France; ³HEMATIM, EA 4666, Université Picardie Jules Verne, Amiens, France; ⁴Assistance Publique-Hôpitaux de Paris, Centre-Université Paris Cité, Service d'Hématologie Biologique, Hôpital Cochin, Paris, France; ⁵Service d'Hématologie-Immunologie-Transfusion, CHU Ambroise Paré, INSERM UMR 1184, AP-HP, Université Paris Saclay, 92100 Boulogne Billancourt, France; ⁶Broad Institute of MIT and Harvard, Cambridge, MA, USA and ⁷Division of Hematology/Oncology, Boston Children's Hospital, Harvard Medical School, Cambridge, MA, USA

Correspondence: T. Boyer
boyer.thomas@chu-amiens.fr

Received: November 3, 2022.

Accepted: March 7, 2023.

Early view: March 16, 2023.

<https://doi.org/10.3324/haematol.2022.282370>

©2023 Ferrata Storti Foundation

Published under a CC BY-NC license



Abstract

The diagnosis of myelodysplastic syndromes (MDS) might be challenging and relies on the convergence of cytological, cytogenetic, and molecular factors. Multiparametric flow cytometry (MFC) helps diagnose MDS, especially when other features do not contribute to the decision-making process, but its usefulness remains underestimated, mostly due to a lack of standardization of cytometers. We present here an innovative model integrating artificial intelligence (AI) with MFC to improve the diagnosis and the classification of MDS. We develop a machine learning model through an elasticnet algorithm directed on a cohort of 191 patients, only based on flow cytometry parameters selected by the Boruta algorithm, to build a simple but reliable prediction score with five parameters. Our AI-assisted MDS prediction score greatly improves the sensitivity of the Ogata score while keeping an excellent specificity validated on an external cohort of 89 patients with an Area Under the Curve of 0.935. This model allows the diagnosis of both high- and low-risk MDS with 91.8% sensitivity and 92.5% specificity. Interestingly, it highlights a progressive evolution of the score from clonal hematopoiesis of indeterminate potential (CHIP) to high-risk MDS, suggesting a linear evolution between these different stages. By significantly decreasing the overall misclassification of 52% for patients with MDS and of 31.3% for those without MDS ($P=0.02$), our AI-assisted prediction score outperforms the Ogata score and positions itself as a reliable tool to help diagnose MDS.

Introduction

Myelodysplastic syndromes (MDS) are a heterogeneous group of myeloid neoplasms the incidence of which increases with age, with a median age at diagnosis of 75 years.¹ MDS are characterized by ineffective hematopoiesis leading to peripheral cytopenia, and dysplastic features in the erythroid, myeloid, monocytic and megakaryocytic cell lineages in bone marrow (BM) and peripheral blood (PB). Besides the symptoms and complications associated with cytopenia, MDS have a time-dependent heterogeneous but life-threatening potential for malignant transformation into acute myeloid leukemia (AML).²

According to the World Health Organization (WHO) 2016 classification, cytomorphology and cytogenetics are the gold standard for MDS diagnosis.³ However, cytomorphological analysis of BM smears may be challenging and identification of myelodysplastic features requires a well-

trained cytologist. It is especially important when cytogenetic analysis does not reveal any chromosomal abnormality, which happens in 50% of patients. Thus, additional diagnostic procedures such as next generation sequencing (NGS) and multiparametric flow cytometry (MFC) need to be performed to facilitate the diagnosis. Several MDS MFC-scores have already been reported, like the Ogata score⁴ focusing on progenitor cells, the RED-score⁵ analyzing nucleated red blood cells, or the integrated flow score (iFS)⁶ that includes aspects of most of the MFC scores. However, the widespread use of these diagnostic tools is severely limited by a lack of standardization of the procedure among the centers performing flow cytometry, especially when using different flow cytometry machines and different brands of antibodies.

Over the past few years, due to the increased sensitivity of molecular biology techniques, several definitions and classifications of pre-MDS conditions have been pro-

posed.⁷ Among these pre-MDS conditions are ICUS (idiopathic cytopenia of unknown significance), CHIP (clonal hematopoiesis of indeterminate potential), and CCUS (clonal cytopenia of unknown significance). They have relied either on the presence (ICUS and CCUS) or the absence (CHIP) of cytopenia, or on the detection (CHIP and CCUS) or not (ICUS) of clonal hematopoiesis in molecular biology or cytogenetic analyses. These stages do not require any treatment and CHIP is associated with a 1% risk of transformation to MDS per year. The differential diagnosis between these pre-MDS stages may be challenging and have implications on the patient's follow-up. Besides, it is often difficult to distinguish pre-MDS stages from low-risk MDS in the absence of marked BM dysplasia, whereas there is a huge difference in the risk of ultimate malignant transformation.⁸

Recently, artificial intelligence (AI) has begun to play an important role in numerous areas of medicine. Several methods (machine learning, convolutional neural networks) have been developed in hematology to address these specific problems. Interestingly, more and more feature selection algorithms are being developed, which allow us to select and focus on the most important parameters of a given pathology.

In this study, using a 10-color single-tube, we sought to discriminate patients with MDS from patients without MDS based on the profile obtained with MFC in a well-characterized and multicentric cohort from hematology departments of three different centers. After features selection using the Boruta algorithm, we established a diagnostic score to accurately distinguish between patients without MDS and patients with or without excess blasts.

Methods

Patients

Between 2019 and 2021, patients with suspected MDS who had undergone MFC evaluation at initial diagnosis were retrospectively identified in the hematology departments of 3 different centers (Amiens, Ambroise Paré and Cochin hospitals). Peripheral blood (PB) cytopenia were defined as platelets below $150 \times 10^9/L$ (thrombocytopenia), neutrophils below $1.8 \times 10^9/L$ (neutropenia), and hemoglobin concentration below 12 g/dL or 13 g/dL (anemia) for women and men, respectively. All MDS diagnoses were made according to the 2016 World Health Organization (WHO) classification. Clinical, morphologic, immunophenotypic, molecular, and cytogenetic data were reviewed. The Revised International Prognostic Scoring System (IPSS-R) was calculated as previously described.⁹ We divided the total cohort into two; first, we designed a learning cohort with patients from two hospitals (Ambroise Paré and Cochin), and then we used an external validation

cohort with patients from the Amiens hospital. This process is the gold standard of development and medical application on a machine learning algorithm. This method, associated with crossvalidation on a learning cohort, allows generalized performances which can be applied to other hospitals, as well as algorithms with less overfitting to be obtained.¹⁰

Flow cytometry

One center used a Navios™ instrument (Beckman Coulter, Miami, FL, USA) and the two others used FACSCanto™ and FACSLyric™ instruments (Beckton Dickinson, Franklin Lakes, NJ, USA). Before each series, the settings of the photomultipliers were checked with fluorescent calibration beads.

Direct immunolabeling was performed on 50 μL of whole BM. After 20 minutes incubation, red blood cells were lysed with Versalysse™ solution (Beckman Coulter) according to the manufacturer's instructions, and the samples were washed once in phosphate-buffered saline (PBS). At least 50,000 events were acquired. All the antibodies used are listed in *Online Supplementary Tables S1-S3*. As described by Della Porta *et al.*¹¹ regarding the Ogata score, four parameters (1 point each when outside the normal ranges) were analyzed: 1) the percentage of CD34⁺ myeloid progenitors among all acquired cells (threshold for normal <2%); 2) the percentage of B-cell progenitors, defined as CD34⁺CD38⁺CD19⁺ cells, among all CD34⁺ cells (threshold for normal >5%); 3) the lymphocyte/myeloid progenitor CD45 ratio (normal range 4-7.5); and 4) the granulocyte/lymphocyte SSC peak channel ratio (threshold for normal >6). The Ogata score was positive if ≥ 2 .

Expression of CD7, CD56 and HLA-DR on blast cells was also assessed, as well as the total hematogone ratio (number of hematogones/number of CD34⁺ cells).

Statistical analysis

R software 4.0.5 was used for the statistical analysis: χ^2 test for categorical variables, non-parametric Kruskal Wallis test and Pearson correlation for quantitative parameters. We performed features selection using the Boruta algorithm with 150 random forest iterations and obtained a predictive model by logistic regression penalized by an elasticnet algorithm on previously selected features.¹² We used an α coefficient of penalization of 0.6 and a 10-fold crossvalidation on training and test cohorts to reduce bias and optimize threshold category; an algorithm performance with a matrix of prediction on the Amiens Hospital validation cohort was obtained. A Receiver Operating Characteristic (ROC) curve analysis was performed on this validation cohort. Finally, we used a Cochran-Mantel-Haenszel χ^2 test to analyze differences between the different matrices of prediction obtained by the Ogata and the elasticnet score on each group.

Machine learning was performed on MDS and no MDS patients without CHIP (n=280). CHIP was added to the figures a posteriori to observe the behavior of the model on this pathology. Two-tailed $P < 0.05$ was considered statistically significant.

Results

A total of 105 (34.65%) patients without MDS (no MDS), 23 with pre-MDS (7.59%), 112 (36.96%) with MDS without excess blasts, and 63 patients (20.79%) with MDS with excess blasts (MDS-EB) were enrolled in the total cohort (Table 1). All patients with pre-MDS stages listed in the introduction were combined in the pre-MDS group because of the small number of patients in each category. Among the 105 no MDS cases, there are patients with drug toxicity (n=15, 14.28%), autoimmune disease (n=10, 9.52%), liver insufficiency (n=12, 11.43%), bone marrow metastasis (n=8, 7.61%), idiopathic thrombocytopenia (ITP) (n=12, 11.43%), infections (n=12, 11.43%), vitamin B9 or B12 deficiency (n=7, 6.67%), non-Hodgkin lymphoma (n=12, 11.43%), aplastic anemia (n=7, 6.67%), kidney failure (n=6, 5.71%), and multiple myeloma (n=4, 3.81%). In the total cohort, there was a significant difference in

median age at diagnosis between these groups (72 years for patients without malignancy, 75 for pre-MDS, 78 for MDS, and 80 for MDS-EB; $P=0.016$). We did not find any significant difference between MCV values, probably due to the high values in some no MDS patients (vitamin deficiency). A positive-Ogata score classified 71% of MDS patients in the MDS group, and accurately classified 81.10% of no MDS patients. It performed better for patients with MDS-EB, accurately classifying 95% of them in the MDS group, whereas it performed less well for the diagnosis of MDS without excess of blasts, both for MDS-MLD and MDS-SLD, only identifying 57.10% and 47.13% of them, respectively. Thus, the sensitivity of the Ogata score was 69.80% with 93.80% specificity, with a 95% positive predictive value (PPV) and a 63% negative predictive value (NPV). As expected, the percentage of CD34⁺ myeloid progenitors was significantly higher in the MDS-EB group (2.76%) than in the patients with no MDS (0.71%), pre-MDS (0.63%), and MDS without excess of blasts (0.92%) ($P < 0.001$). We found no significant immunophenotypic aberrations, with a similar median value of CD7⁺ blasts ($P=0.942$), CD56⁺ blasts ($P=0.551$), and HLA-DR negative blasts ($P=0.658$) across the different groups. On the other hand, the percentage of CD7⁺ blast cells was significantly higher in the MDS groups ($P < 0.001$).

Table 1. Global cohort characteristics grouped by pathology.

Parameters Median (range)	Pathology				P
	No MDS N=105	CHIP N=23	MDS N=112	MDS-EB N=63	
Age in years	72 (33-92)	75 (43-88)	78 (45-92)	80 (54-91)	0.016
Hemoglobin, g/dL	11.4 (5.4-17.7)	10.8 (9-15.8)	10.6 (6.1-16.3)	9.5 (7.7-13)	0.059
MCV, fL	93 (64.3-110)	95.9 (74-109)	98 (79.8-119)	95.6 (83-125)	0.212
Platelet count, x10 ⁹ /L	111 (11-479)	136 (35-918)	146 (14-608)	82 (15-320)	0.048
WBC, x10 ⁹ /L	3.6 (1.1-12.8)	6.2 (2.2-24)	4.6 (1.2-12.2)	3.2 (1.5-7.8)	0.005
ANC, x10 ⁹ /L	2.2 (0.3-8.7)	4.1 (0.9-20)	2.9 (0.59-10.2)	1.9 (0.3-6.4)	0.01
Monocyte count, %	0.4 (0.07-18)	1.5 (0.3-11.2)	0.54 (0-14)	0.29 (0-4.3)	0.001
Bone marrow blasts, %	2 (0-4)	2 (2-3)	2 (0-4)	8 (5-18)	<0.001
CD34 ⁺ myeloid progenitors, %	0.7 (0.06-3.5)	0.6 (0.05-2.2)	0.9 (0.08-5.2)	2.76 (0.5-23)	<0.001
CD34 ⁺ CD38 ⁻ , %	6.3 (0.1-36)	11 (0.3-21)	6.75 (0-39)	4.54 (0.25-60)	0.012
CD7 ⁺ blasts, %	6.33 (0-25.6)	6 (1.2-38.5)	7 (0-82)	7 (0-87)	0.94
CD56 ⁺ blasts, %	0.8 (0-9.1)	0.4 (0-16.8)	1 (0-60.6)	1.2 (0-85)	0.55
HLA-DR ⁻ blasts, %	1.6 (0-30)	1.6 (0.2-10.8)	1.65 (0-56.5)	2.6 (0.5-13)	0.66
Total Hg, ratio	134.8 (1-1,172)	99.2 (11.3-539)	21.7 (0.3-647)	3.4 (0-114.8)	<0.001
Hg Ogata, %	19.5 (0-89.2)	11.5 (0.6-48.7)	3.94 (0-67.6)	0.8 (0-37)	<0.001
CD45 Ogata, ratio	6.8 (3.23-20.7)	11.7 (9.1-16.8)	6.7 (4.4-23.1)	8 (3.5-34.7)	0.4
SSC Ogata, ratio	6.95 (4.2-15.3)	5.8 (3.8-7.6)	5.5 (3.3-15.6)	5.1 (2.1-12.6)	<0.001

MDS: myelodysplastic syndromes; CHIP: clonal hematopoiesis of indeterminate potential; MDS-EB: MDS with excess blasts; MCV: mean corpuscular volume; WBC: white blood cells; ANC: absolute neutrophil count; CD34⁺CD38⁻: percentage of CD34⁺ CD38⁻ blast cells; CD7⁺ blasts: median value of CD7⁺ blast cells; CD56⁺ blasts: median value of CD56⁺ blast cells; HLA-DR⁻ blasts: median value of HLA-DR⁻ blast cells; Total Hg: ratio Hg/CD34⁺; Hg Ogata: hematogone parameter of the Ogata Score; CD45 Ogata score: CD45 lymphocytes/ myeloid progenitor CD45 ratio of the Ogata score; SSC Ogata: granulocytes/lymphocytes side scatter parameter of the Ogata score.

To help distinguish patients with MDS from those with no MDS, we applied a Boruta feature selection algorithm on flow cytometry parameters. The qualitative granulocyte/lymphocyte SSC peak channel ratio (SSC Ogata score <6), the total hematogone ratio (number of hematogones/number of CD34⁺ cells), the percentage of CD34⁺ B-cell progenitors among all CD34⁺ cells (hematogone Ogata score), and the percentage of CD34⁺ myeloid progenitors were informative features to predict the diagnosis of MDS by MFC. On the contrary, the percentage of CD34⁺CD38⁻ blast cells and the lymphocyte/myeloid progenitor CD45 ratio were found to be the least relevant features (Figure 1).

We then used an elasticnet model to evaluate the probability of MDS (with or without excess blasts) on the learning and the test cohorts (plus pre-MDS patients; Figure 2A), allowing us to propose this mathematical formula: MDS prediction score = -1.58 + 2.928*SSC Ogata score + 0.965*hematogone Ogata score + 0.8907*%CD34⁺ myeloid progenitors - 0.0032*Total hematogone ratio. In our learning cohort, a threshold of 0 with this formula proved ideal to distinguish patients with MDS from pre-MDS or no MDS patients using crossvalidation. We then validated this model with a threshold of 0 on the Amiens Hospital validation cohort, and we obtained a ma-

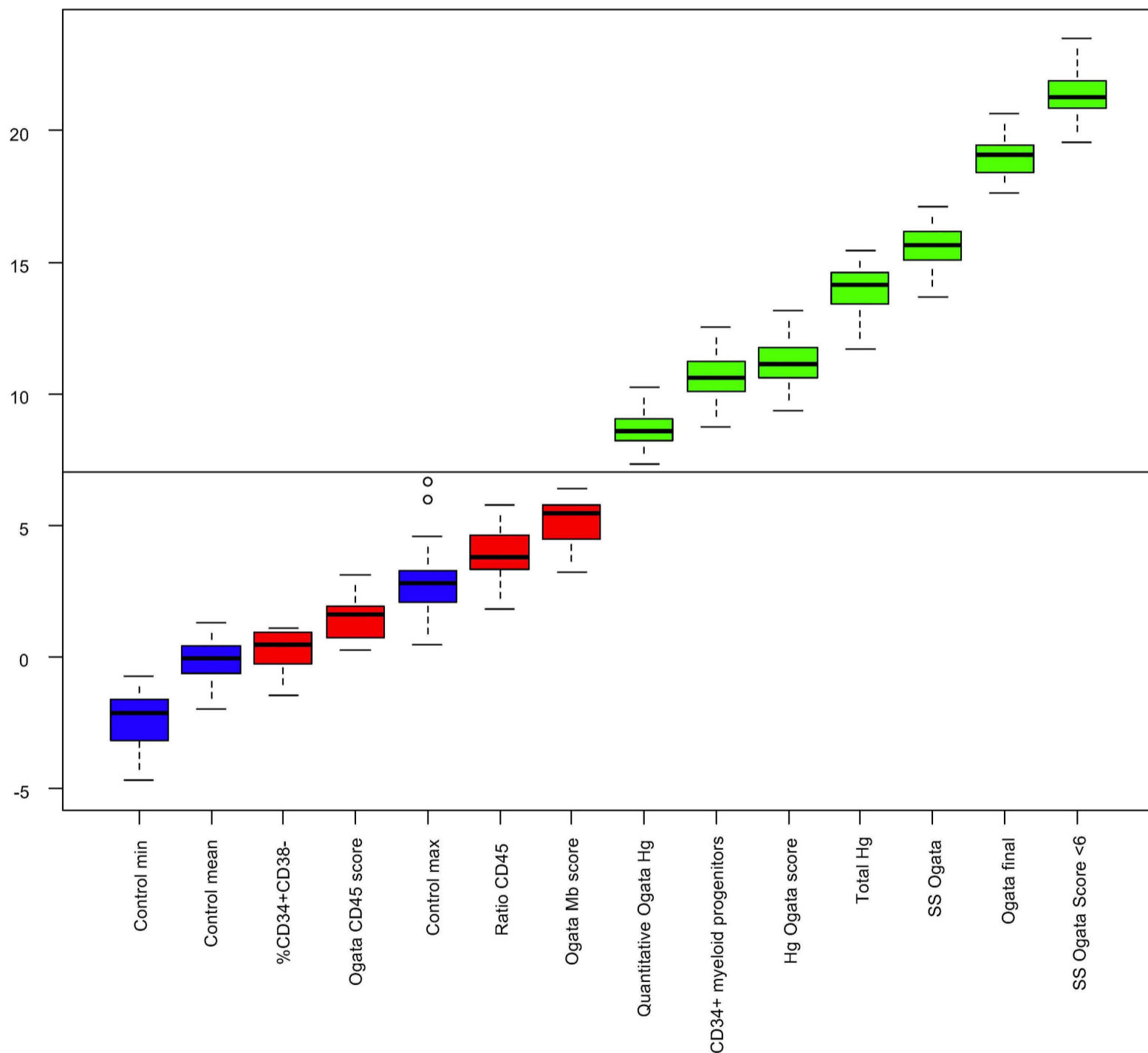


Figure 1. Features selection by Boruta algorithm. Features in green are validated as useful, features in red are useless to predict myelodysplastic syndromes, control in blue are hazard permuted features that allow us to confront predictive potential of our variables with these randomness features. y-axis represents difference between each Z-score accuracy on feature and control, associated with standard deviation obtained on 100 iterations. Horizontal threshold represented significant different threshold between features and control. SS Ogata score: granulocyte/lymphocyte SSC peak channel ratio; Ogata final: numerical value of the Ogata score between 0 and 4; SS Ogata: numerical value of the granulocyte/lymphocyte SSC peak channel ratio; Total Hg: number of hematogones/number of CD34⁺ cells; Hg Ogata score: percentage of CD34⁺ B-cell progenitors among all CD34⁺ cells; % of MFC blasts: percentage of CD34⁺ myeloid progenitors; quantitative Ogata Hg: numerical value of the percentage of CD34⁺ B-cell progenitors among all CD34⁺ cells; Ogata Mb score: percentage of CD34⁺ myeloid progenitors; Ratio CD45: numerical value of the lymphocyte/myeloid progenitor CD45 ratio; Ogata CD45 score: out of 4-7.8 range; min: minimum; max: maximum.

trix prediction with sensitivity of 91.84%, a specificity of 92.48%, and positive and negative predictive values of 93.75% and 90.03%, respectively. The prognostic values of our AI-assisted MDS prediction score strikingly outperforms the Ogata score, particularly by significantly improving the sensitivity and the NPV (Table 2).

To validate our diagnosis threshold with normalized data, we built a cumulative proportion plot for all pathologies on the total cohort that shows representativeness of this cohort without imbalanced data (Figure 2B).

With a cut off of 0, ROC curve analysis performed on the same validation cohort showed an Area Under the Curve of 0.935, highlighting the excellent performance of the MDS prediction score by which only 6.47% of the patients were misclassified (Figure 2C).

In the total cohort (191 patients on learning and test cohorts plus 89 patients on external validation cohort: total $n=280$), our elasticnet MDS prediction score allowed for a clear distribution of MDS, MDS-EB and no MDS groups (Figure 2D). To improve the accuracy of diagnosis for these three subgroups, we refined the thresholds of our model

and identified three groups: group A with an elasticnet MDS prediction score between -3 and 0 , group B higher than 0 and less than 3 , and group C higher than 3 (Figure 3). Group A included many patients without MDS (87.88% with no MDS; 90.90% of the no MDS patients in the global cohort were in this group: $n=132$). Group B had more MDS patients without EB (88.17% with MDS in this group: $n=93$), and group C had more MDS-EB patients or MDS without EB but with multiple abnormalities like multi-lineage dysplasia or genetic abnormalities (100% of MDS in this group: $n=78$) (Table 3). Patients with pre-MDS stages were equally distributed in group A and group B. Strikingly, our AI-assisted model shows a progressive evolution of the MDS prediction score from a pre-MDS condition (CHIP) to high-risk MDS, suggesting a linear evolution between these different stages (Figure 2C and Figure 4).

We then tested the accuracy of the model to classify patients according to the IPSS-R categories. In the whole cohort, IPSS-R was available for 150 MDS patients. Only 18 MDS patients (9.42% of total MDS; 13.63% of group A) were classified in group A, all others being classified in groups

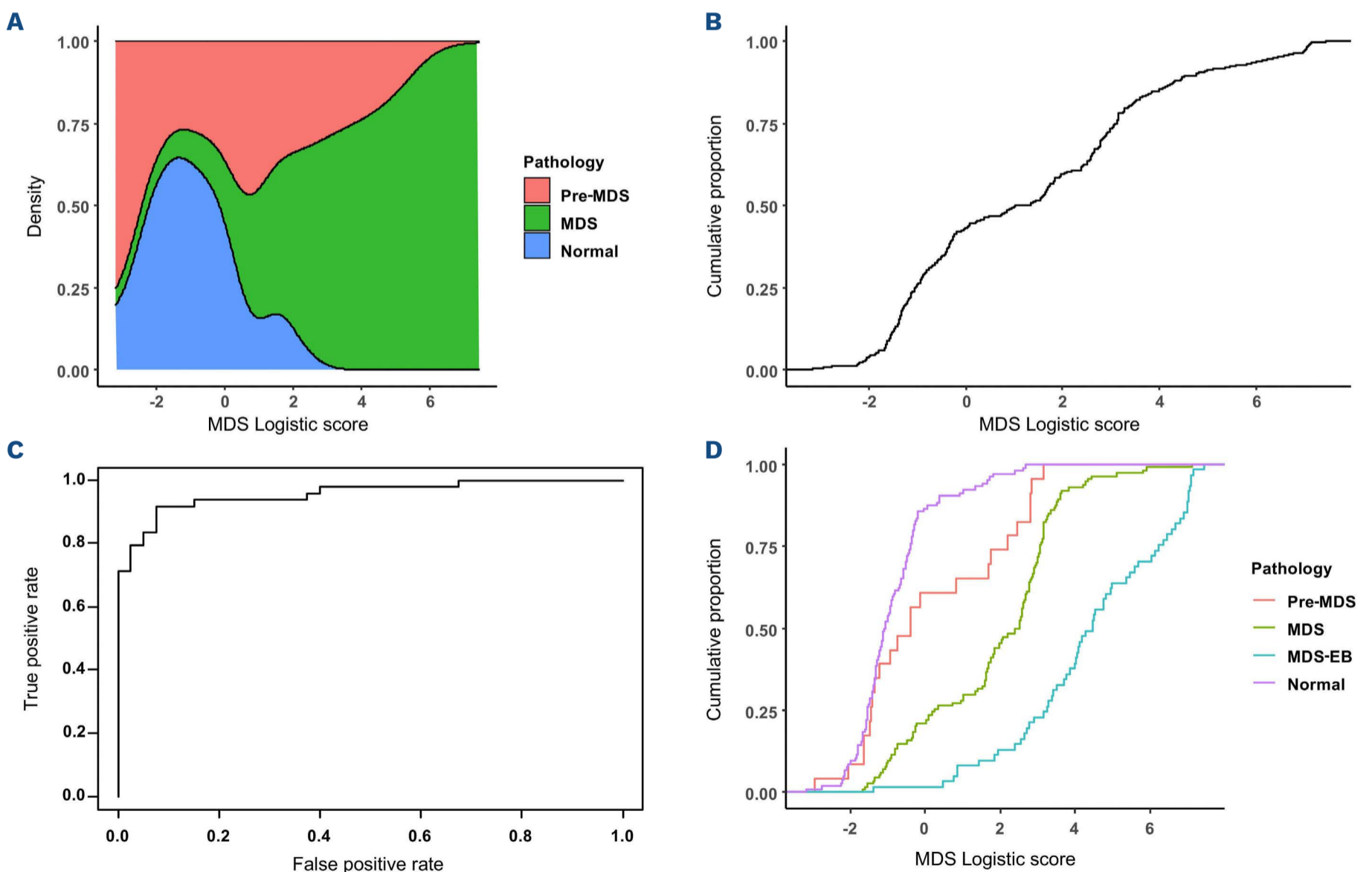


Figure 2. Probability of myelodysplastic syndrome (MDS) diagnosis according to the elasticnet MDS score. Development of an elasticnet model to predict the probability of MDS in the training and test cohorts. (A) Density plot of pathology repartition according to the MDS score. (B) Cumulative proportion of the cases in the global cohort according to the MDS score. (C) Receiver Operating Characteristic curves of the MDS score on the external validation cohort for MDS and no MDS patients (without pre-MDS patients). (D) Cumulative proportion of the MDS score according to different pathologies. Pre-MDS: pre-myelodysplastic syndrome conditions; MDS-EB: MDS with excess blasts.

Table 2. Matrix of elasticnet scoring prediction algorithm on the Amiens Hospital validation cohort.

Score-based prediction	Pathology			Performance			
	MDS N=49	No MDS N=40	ER	Sensibility	Specificity	PPV	NPV
MDS	45	3	6.25%	91.8%	92.5%	93.75%	90%
No MDS	4	37	9.7%	-	-	-	-
ER	8%	7.5%	8%	-	-	-	-

MDS: myelodysplastic syndromes; ER: error rate based on the misclassified rate; PPV: positive predictive value; NPV: negative predictive value.

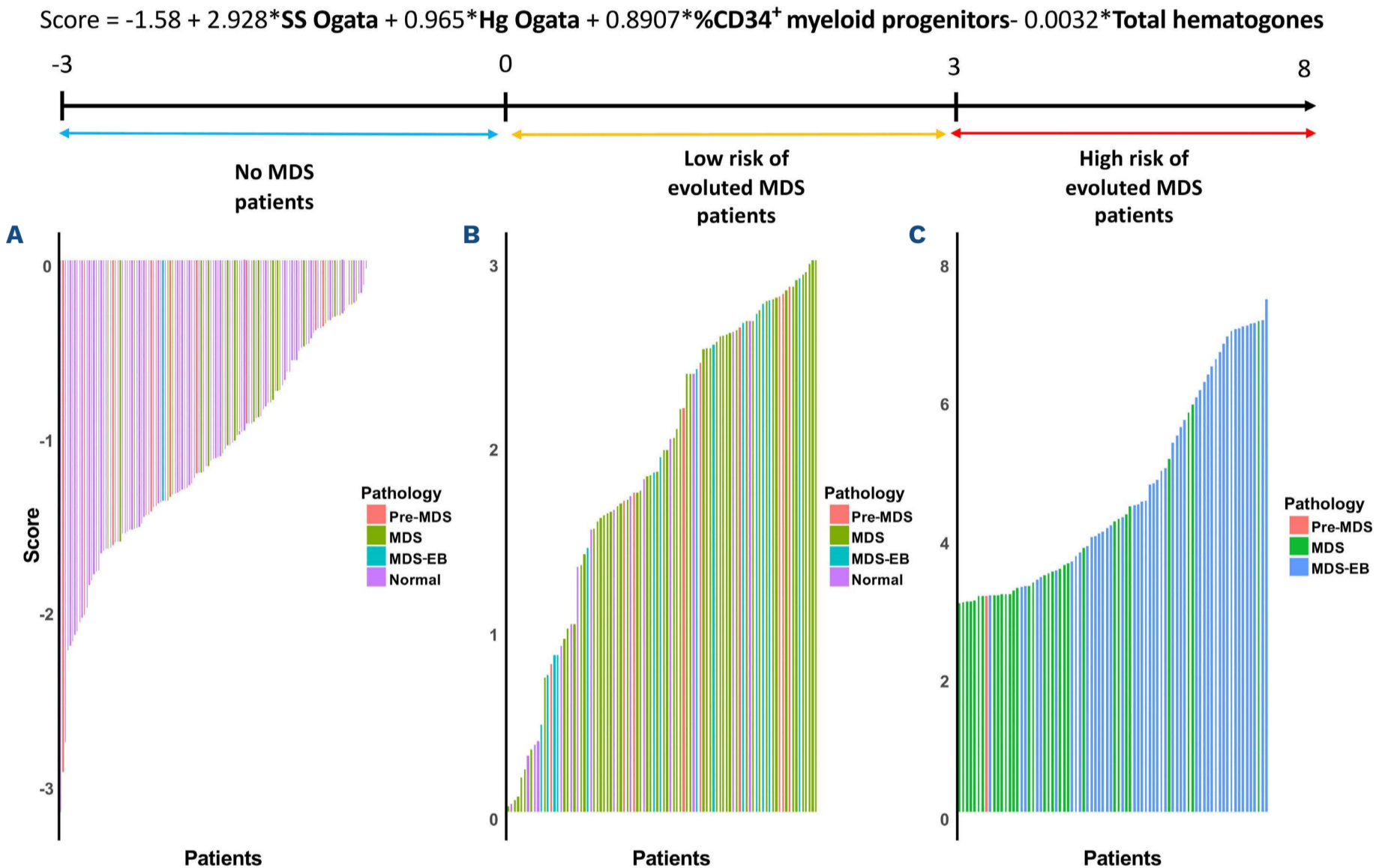


Figure 3. Elasticnet myelodysplastic syndrome model and distribution of diseases in the three risk groups. Proposed score scale with corresponding group breakdowns: (A) between -3 and 0, (B) between 0 and 3, and (C) between 3 and 8. MDS-EB: myelodysplastic syndrome with excess blasts; Pre-MDS: pre-myelodysplastic syndrome conditions; MDS: myelodysplastic syndromes (with single or multilineage dysplasia).

B and C of our model. Of these 18 patients, only one was of intermediate risk according to the IPSS-R, whereas nine were in the low-risk group and eight in the very low-risk group. Importantly, no high-risk patients were classified in group A (Table 4).

Overall, our newly developed AI-assisted MDS prediction score improved the accuracy of MDS diagnosis, by reducing the risk of misclassification of MDS without excess blasts by 52.07% ($P=0.004$) and an absence of MDS by 31.33% ($P=0.022$) compared to the Ogata score (Figure 4). Therefore, the sensitivity of this score for the subgroup of

patients without excess blasts was 78.27%. These different subtypes could not be diagnosed by flow cytometry (using cytology alone) but were informative as to the ability of the algorithm to predict MDS.

Discussion

In this study, we developed an innovative model combining artificial intelligence and machine learning with flow cytometry to improve the performance of MFC in diagnosing MDS.

Table 3. Parameters grouped by logistic model.

Parameters Median (range)	Logistic model			P
	No MDS group N=132	MDS low-risk group N=93	MDS high-risk group N=78	
Age in years	72 (33-92)	78 (39-92)	78 (54-91)	0.13
Hemoglobin, g/dL	10.9 (6.2-17.7)	10.5 (6.6-16.3)	9.8 (6.2-15.8)	0.03
MCV, fL	95.8 (73.1-119)	95.6 (64.3-115)	95.7 (83-124)	0.55
Platelet count, x10 ⁹ /L	120 (11-568)	136 (14-918)	114 (14-617)	0.65
WBC, x10 ⁹ /L	4.2 (1.3-14.8)	5.4 (1.1-24)	3.8 (1.2-10)	0.12
ANC, x10 ⁹ /L	2.7 (0.5-12.7)	3.2 (0.3-23.2)	2.1 (0.3-6.4)	0.07
Monocyte count, %	0.4 (0.1-18)	0.5 (0.1-22)	0.4 (0-4.3)	0.09
Bone marrow blasts, %	2 (0-6)	2 (0-5)	3 (1-18)	0.01
CD34 ⁺ myeloid progenitors, %	0.7 (0.05-2.2)	0.7 (0.08-3.89)	2.5 (0.9-16.5)	<0.001
CD34 ⁺ CD38 ⁻ , %	6.4 (0.1-36)	7.3 (0-26.2)	5.3 (0.24-77)	0.38
CD7 ⁺ blasts, %	7.3 (0-38.5)	5.8 (0.03-82)	7.8 (0-87)	0.41
CD56 ⁺ blasts, %	0.8 (0-16.8)	0.9 (0-60.7)	1 (0-85)	0.73
HLA-DR ⁻ blasts, %	1.7 (0-30)	1.4 (0-13.3)	2.3 (0.1-77)	0.11
Total Hg, ratio	94.8 (0.99-1,172)	23.9 (0.22-465)	3.6 (0-96)	<0.001
Hg Ogata, %	20.1 (0-92.4)	4.5 (0-87.6)	0.9 (0-13.9)	<0.001
CD45 Ogata, ratio	6.8 (3.77-20.7)	7.4 (3.23-48)	8.2 (3.5-34.7)	0.35
SS Ogata, ratio	7.1 (4.2-15.3)	5.5 (3.3-13.82)	4.9 (2.1-9.4)	<0.001

MDS: myelodysplastic syndromes; MCV: mean corpuscular volume; WBC: white blood cells; ANC: absolute neutrophil count; CD34⁺CD38⁻: percentage of CD34⁺ CD38⁻ blast cells; CD7⁺ blasts: median value of CD7⁺ blast cells; CD56⁺ blasts: median value of CD56⁺ blast cells; HLA-DR⁻ blasts: median value of HLA-DR⁻ blast cells; Total Hg: ratio Hg/CD34⁺; Hg Ogata: hematogone parameter of the Ogata score; CD45 Ogata: CD45 lymphocytes/myeloid progenitor CD45 ratio of the Ogata score; SSC Ogata: granulocytes/lymphocytes side-scatter parameter of the Ogata score.

Here, we propose an original AI-assisted prediction score for MDS to directly investigate the value of combining AI and MFC for the diagnosis of MDS. A few studies have used convolutional neural networks with gradient boosting to assess dysplasia^{13,14} or to distinguish aplastic anemia from MDS with very good sensitivity and specificity;¹⁵ but, up till now, AI has remained underused in the diagnosis of MDS, particularly in combination with MFC. Two other studies demonstrated a link between morphology, mutational status and prognosis in MDS using machine learning techniques.^{16,17} Recent studies used unsupervised cluster analysis of flow cytometry to identify new subsets in pathological erythropoiesis or facilitate the diagnosis of MDS.^{18,19}

The aim of this study was to distinguish patients with actual MDS from patients without MDS. Our AI-assisted MDS prediction score following an elasticnet model identified that most of the MDS patients have a score >0. By setting the cut-off value to 0, our diagnostic model shows high predictive value and strikingly outperforms the Ogata score by significantly increasing the sensitivity and the specificity of this test. If the original Ogata score performs well to discriminate MDS-EB (that are also usually easier to diagnose on BM smears), the great value of our score is in improving the accuracy of diagnosing MDS without excess blasts, whether they show single or multi-lineage

dysplasia. With an error rate of 8% for both false positive and negative results, our MDS prediction score will increase user confidence for biologists and clinicians involved in the diagnostic procedure, especially when the presence of dysplastic features is not clear. In addition to its performances in the diagnosis of MDS, our model allowed for risk prediction, as we identified three risk groups (A, B and C) that correlate with the evolution of the disease. Nevertheless, one MDS-EB patient was classified in group A (see Figure 3.) This patient had bicytopenia, 6% blasts on the bone marrow aspiration with no cytogenetic abnormalities and no mutations found, and was classified in the IPSS-R category. The patient is currently being monitored and is not receiving treatment. This is an unusual case, and the excess of blasts on bone marrow smears could be reactive (this excess was not found on MFC analysis with an Ogata score = 0). The follow-up of this patient could help us to understand his classification in group A.

In our model, patients with pre-MDS cluster equally with patients without MDS. This specific distribution argues for a continuity between the occurrence of clonal hematopoiesis and the onset of MDS, as previously suggested by several teams.^{7,20-25}

Our prediction score included a few patients diagnosed with ICUS and CCUS; these had to be combined with CHIP

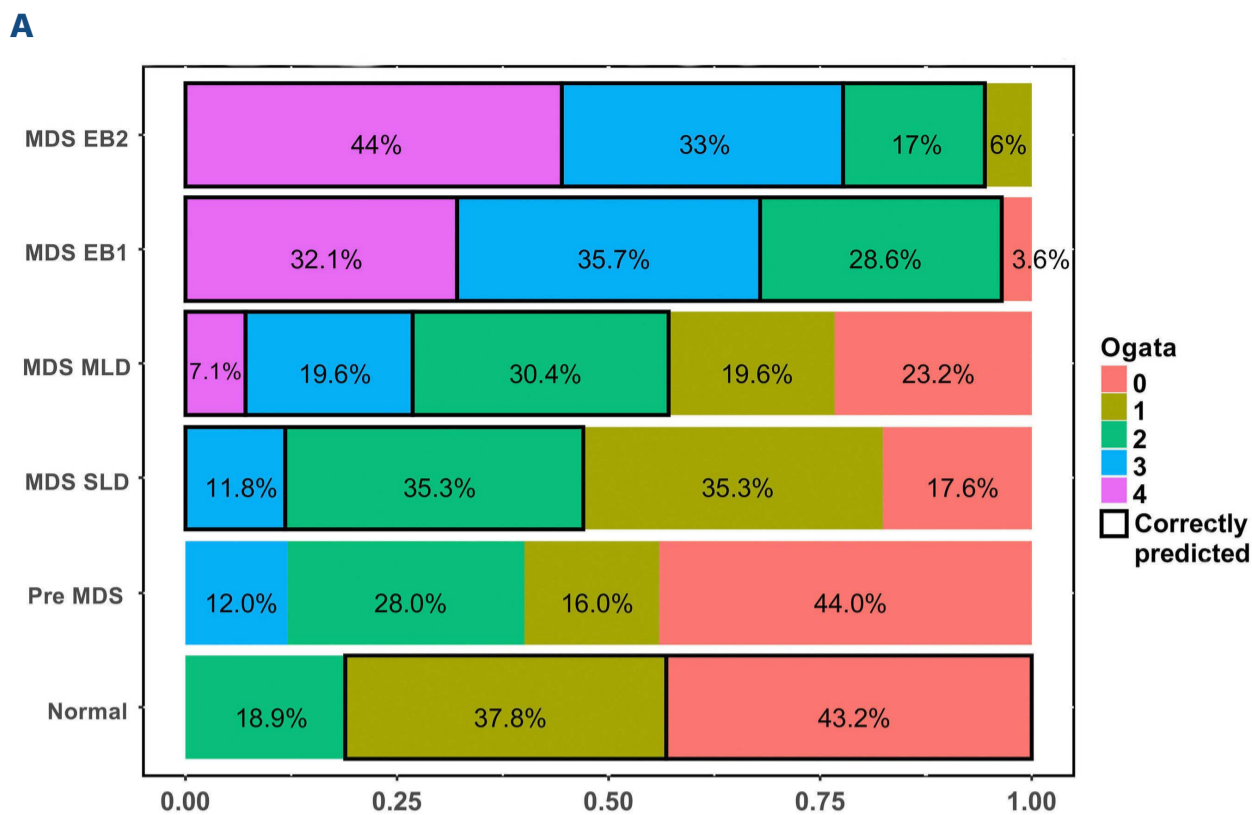
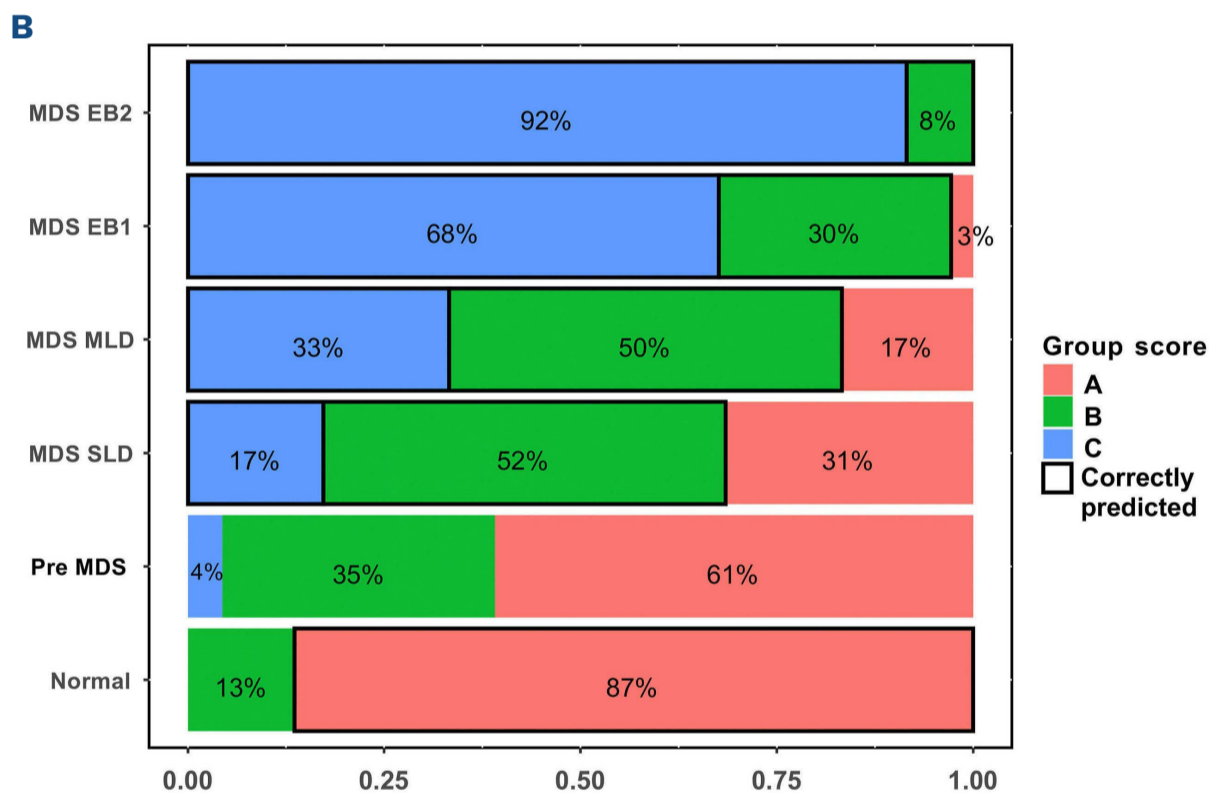


Figure 4. Distribution of myelodysplastic syndrome (MDS) prediction on the different groups with the elasticnet MDS score and Ogata score.

Distribution of the different subtypes of pathologies according to the Ogata score and according to the penalized logistic score group. Diagnosis of myelodysplastic syndrome with single lineage dysplasia (MDS-SLD) and myelodysplastic syndrome with multilineage dysplasia (MDS-MLD) was made in cytology. (A) Classification according to the Ogata score. (B) Classification according to the elasticnet score. MDS-EB2: myelodysplastic syndrome with excess blasts between 10% and 19%; MDS-EB1: myelodysplastic syndrome with excess blasts between 5% and 9%; pre-MDS: pre-myelodysplastic syndrome conditions.



patients to preserve the score's performance. It would be interesting to further explore their behavior in our model. Furthermore, we could not discriminate between the different categories of MDS (e.g., 5q- syndrome, MDS with ring sideroblasts), and this could be the subject of further tests. It would also be interesting to know whether a positive score increases the risk of developing an overt MDS. Unfortunately, the cohort was not designed to answer this question, and more patients and a longer follow-up are required.

The model has been built using the 2016 MDS WHO classification. As the recent release of the new classification might change the performances values, we aimed to reclassify our patients according to the new WHO classification.

As no patient was diagnosed with MDS-EB2 with AML-defining cytogenetics, this update did not change the initial classification of the patients.

Another potential limitation of the generalizability of the model is the lack of standardization between the different centers when assessing the Ogata score. Despite published recommendations,²⁶ centers either use different fluorochromes or different clones of antibodies. However, when using data with artificial intelligence, this variability is a real asset in helping to build the algorithm. Moreover, all of the three centers involved in the study passed the external quality controls for this panel of antibodies on normal BM samples through the Acute Leukemia French Association / French Innovative Leukemia Organization

(ALFA-FILO) network. Despite using different antibodies and dyes, the MDS prediction score yielded excellent results in the validation cohort, suggesting our model could be widely used.

Flow cytometry provides faster results than most cytogenetics or molecular biology techniques and is widely available worldwide; its standardization between laboratories is, therefore, of crucial importance. It relies on the same panel as that used in the Ogata score, which is already carried out in most laboratories. At a time when cost-effectiveness is becoming increasingly important, this AI-assisted MDS prediction score enables rapid patient diagnosis and stratification to help clinicians in their quest for the best patient care.

Disclosures

No conflicts of interest to disclose.

Contributions

TB and VC designed the research study. TB, NC, JZ, VB and LG collected and analyzed the data. DL, AC and JPM managed patients and provided clinical data. TB, VC, VB, NC, LG and AC wrote the paper. All authors approved the final version of the manuscript for publication.

Data-sharing statement

The authors confirm that the data supporting the findings of this study are available within the article and its Online Supplementary Appendix.

References

1. Zeidan AM, Shallis RM, Wang R, Davidoff A, Ma X. Epidemiology of myelodysplastic syndromes: why characterizing the beast is a prerequisite to taming it. *Blood Rev.* 2019;34:1-15.
2. Menssen AJ, Walter MJ. Genetics of progression from MDS to secondary leukemia. *Blood.* 2020;136(1):50-60.
3. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood.* 2016;127(20):2391-2405.
4. Ogata K, Della Porta MG, Malcovati L, et al. Diagnostic utility of flow cytometry in low-grade myelodysplastic syndromes: a prospective validation study. *Haematologica.* 2009;94(8):1066-1074.
5. Mathis S, Chapuis N, Debord C, et al. Flow cytometric detection of dyserythropoiesis: a sensitive and powerful diagnostic tool for myelodysplastic syndromes. *Leukemia.* 2013;27(10):1981-1987.
6. Cremers HR, Wager TD, Yarkoni T. The relation between statistical power and inference in fMRI. Gilbert S, ed. *PLoS One.* 2017;12(11):e0184923.
7. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood.* 2015;126(1):9-16.
8. Sekeres MA, Taylor J. Diagnosis and treatment of myelodysplastic syndromes: a review. *JAMA.* 2022;328(9):872-880.
9. Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood.* 2012;120(12):2454-2465.
10. Andaur Navarro CL, Damen JAA, Takada T, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. *BMJ.* 2021;375:n2281.
11. Della Porta MG, Picone C, Pascutto C, et al. Multicenter validation of a reproducible flow cytometric score for the diagnosis of low-grade myelodysplastic syndromes: results of a European LeukemiaNET study. *Haematologica.* 2012;97(8):1209-1217.
12. Kursa MB, Jankowski A, Rudnicki WR. Boruta – a system for feature selection. *Fundam Inform.* 2010;101(4):271-285.
13. Mori J, Kaji S, Kawai H, et al. Assessment of dysplasia in bone marrow smear with convolutional neural network. *Sci Rep.* 2020;10(1):14734.
14. Acevedo A, Merino A, Boldú L, Molina Á, Alférez S, Rodellar J. A new convolutional neural network predictive model for the automatic recognition of hypogranulated neutrophils in myelodysplastic syndromes. *Comput Biol Med.* 2021;134:104479.
15. Kimura K, Tabe Y, Ai T, et al. A novel automated image analysis system using deep convolutional neural networks can assist to differentiate MDS and AA. *Sci Rep.* 2019;9(1):13385.
16. Nagata Y, Zhao R, Awada H, et al. Machine learning demonstrates that somatic mutations imprint invariant morphologic features in myelodysplastic syndromes. *Blood.* 2020;136(20):2249-2262.
17. Nazha A, Komrokji R, Meggendorfer M, et al. Personalized prediction model to risk stratify patients with myelodysplastic syndromes. *J Clin Oncol.* 2021;39(33):3737-3746.
18. Porwit A, Violidaki D, Axler O, Lacombe F, Ehinger M, Béné MC. Unsupervised cluster analysis and subset characterization of abnormal erythropoiesis using the bioinformatic FLOW-SELF Organizing Maps algorithm. *Cytometry B Clin Cytom.* 2022;102(2):134-142.
19. Duetz C, Van Gassen S, Westers TM, et al. Computational flow cytometry as a diagnostic tool in suspected myelodysplastic syndromes. *Cytometry A.* 2021;99(8):814-824.
20. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med.* 2014;371(26):2477-2487.
21. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med.* 2014;371(26):2488-2498.
22. Desai P, Mencia-Trinchant N, Savenkov O, et al. Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nat Med.* 2018;24(7):1015-1023.
23. Abelson S, Collord G, Ng SWK, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature.* 2018;559(7714):400-404.
24. Takahashi K, Wang F, Kantarjian H, et al. Preleukaemic clonal haemopoiesis and risk of therapy-related myeloid neoplasms: a case-control study. *Lancet Oncol.* 2017;18(1):100-111.
25. Gondek LP, DeZern AE. Assessing clonal haematopoiesis: clinical burdens and benefits of diagnosing myelodysplastic syndrome precursor states. *Lancet Haematol.* 2020;7(1):e73-e81.
26. Westers TW, Ireland R, Kern W, et al. Standardization of flow cytometry in myelodysplastic syndromes: a report from an international consortium and the European LeukemiaNet Working Group. *Leukemia.* 2012;26(7):1730-1741.