

# ELN2017 risk stratification improves outcome prediction when applied to the prospective GIMEMA AML1310 protocol

Francesco Buccisano,<sup>1,\*</sup> Raffaele Palmieri,<sup>1,\*</sup> Alfonso Piciocchi,<sup>2</sup> Valentina Arena,<sup>2</sup> Anna Candoni,<sup>3</sup> Lorella Melillo,<sup>4</sup> Valeria Calafiore,<sup>5</sup> Roberto Cairoli,<sup>6</sup> Paolo de Fabritiis,<sup>7</sup> Gabriella Storti,<sup>8</sup> Prassede Salutati,<sup>9</sup> Francesco Lanza,<sup>10</sup> Giovanni Martinelli,<sup>11,12</sup> Mario Luppi,<sup>13</sup> Saveria Capria,<sup>14</sup> Luca Maurillo,<sup>1</sup> Maria Ilaria Del Principe,<sup>1</sup> Giovangiaco Paterno,<sup>1</sup> Maria Antonietta Irno Consalvo,<sup>1</sup> Tiziana Ottone,<sup>1</sup> Serena Lavoragna,<sup>1</sup> Maria Teresa Voso,<sup>1</sup> Paola Fazi,<sup>2</sup> Marco Vignetti,<sup>2</sup> William Arcese,<sup>1,15</sup> and Adriano Venditti<sup>1</sup>

<sup>1</sup>Hematology, Department of Biomedicine and Prevention, Tor Vergata University, Rome, Italy; <sup>2</sup>GIMEMA Foundation, Rome, Italy; <sup>3</sup>Hematology, Azienda Sanitaria Universitaria Integrata di Udine, Udine, Italy; <sup>4</sup>Fondazione IRCCS Casa Sollievo della Sofferenza, UO di Ematologia, San Giovanni Rotondo (FG), Italy; <sup>5</sup>Ospedale Ferrarotto, Catania, Italy; <sup>6</sup>Ospedale Niguarda Ca Granda, Milan, Italy; <sup>7</sup>Ospedale S. Eugenio, Rome, Italy; <sup>8</sup>Azienda Ospedaliera S.G. Moscati, Avellino, Italy; <sup>9</sup>Azienda USL di Pescara, Pescara, Italy; <sup>10</sup>Ospedale S. Maria delle Croci, Ravenna, Italy; <sup>11</sup>Istituto Tumori della Romagna, Meldola, Italy; <sup>12</sup>Policlinico S. Orsola-Malpighi, Bologna, Italy; <sup>13</sup>Ematologia, Dipartimento di Scienza Mediche e Chirurgiche Materno-Infantili e dell'Adulto, Università degli studi di Modena e Reggio Emilia, Modena, Italy; <sup>14</sup>Dipartimento di Biotecnologie Cellulari ed Ematologia, Ematologia Università degli Studi Sapienza, Rome, Italy; and <sup>15</sup>Rome Transplant Network, Rome, Italy

## Key Points

- Within homogeneous ELN risk categories, different postremission approaches (auto-SCT vs allo-SCT) result in different outcomes.
- Measurable residual disease confirms its role as a driver of transplant allocation for patients placed in the ELN2017-IR category.

The 2017 version of the European LeukemiaNet (ELN) recommendations, by integrating cytogenetics and mutational status of specific genes, divides patients with acute myeloid leukemia into 3 prognostically distinct risk categories: favorable (ELN2017-FR), intermediate (ELN2017-IR), and adverse (ELN2017-AR). We performed a post hoc analysis of the GIMEMA (Gruppo Italiano Malattie EMatologiche dell'Adulto) AML1310 trial to investigate the applicability of the ELN2017 risk stratification to our study population. In this trial, after induction and consolidation, patients in complete remission were to receive an autologous stem cell transplant (auto-SCT) if categorized as favorable risk or an allogeneic stem cell transplant (allo-SCT) if adverse risk. Intermediate-risk patients were to receive auto-SCT or allo-SCT based on the postconsolidation levels of measurable residual disease as measured by using flow cytometry. Risk categorization was originally conducted according to the 2009 National Comprehensive Cancer Network recommendations. Among 500 patients, 445 (89%) were reclassified according to the ELN2017 criteria: ELN2017-FR, 186 (41.8%) of 455; ELN2017-IR, 179 (40.2%) of 445; and ELN2017-AR, 80 (18%) of 455. In 55 patients (11%), ELN2017 was not applicable. Two-year overall survival (OS) was 68.8%, 51.3%, 45.8%, and 42.8% for the ELN2017-FR, ELN2017-IR, ELN2017-not classifiable, and ELN2017-AR groups, respectively ( $P < .001$ ). When comparing the 2 different transplant strategies in each ELN2017 risk category, a significant benefit of auto-SCT over allo-SCT was observed among ELN2017-FR patients (2-year OS of 83.3% vs 66.7%;  $P = .0421$ ). The 2 transplant procedures performed almost equally in the ELN2017-IR group (2-year OS of 73.9% vs 70.8%;  $P = .5552$ ). This post hoc analysis of the GIMEMA AML1310 trial confirms that the ELN2017 classification is able to accurately discriminate patients with different outcomes and who may benefit from different transplant strategies. This trial was registered as EudraCT number 2010-023809-36 and at [www.clinicaltrials.gov](http://www.clinicaltrials.gov) as #NCT01452646.

Submitted 8 July 2021; accepted 5 October 2021; prepublished online on *Blood Advances* First Edition 3 November 2021; final version published online 18 April 2022. DOI 10.1182/bloodadvances.2021005717.

\*F.B. and R.P. contributed equally to this study.

Details regarding the original AML1310 trial collection data methods can be obtained by contacting [adriano.venditti@uniroma2.it](mailto:adriano.venditti@uniroma2.it) (refer to EudraCT #2010-023809-36; [www.clinicaltrials.gov](http://www.clinicaltrials.gov) #NCT01452646).

The full-text version of this article contains a data supplement.

© 2022 by The American Society of Hematology. Licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), permitting only noncommercial, nonderivative use with attribution. All other rights reserved.

## Introduction

In 2010 and then in 2017, a panel of experts, convened on behalf of the European LeukemiaNet (ELN), released and updated recommendations for harmonizing diagnosis, response, and prognostic assessment of acute myeloid leukemia (AML).<sup>1,2</sup> The latest version was substantially influenced by evidence regarding the prognostic significance of the *fms*-related tyrosine kinase 3 (*FLT3*) and nucleophosmin 1 (*NPM1*) gene mutations. Mutations of *FLT3*, both at the tyrosine kinase domain (TKD) and the juxta-membrane level (ITD), are recognized in 20% to 25% of patients with a diagnosis of AML, thus representing the most common lesions found in this disease.<sup>3,4</sup> In particular, the *FLT3*-ITD mutation has historically been associated with a poor prognosis due to an increased relapse rate and dismal overall survival (OS). The ELN2017 recommendations have revisited and expanded this concept by capturing the prognostic meaning of the *FLT3*-ITD allelic ratio<sup>5,6</sup> and of concomitant mutations of *FLT3*-ITD and *NPM1*. Actually, mutations of the *NPM1* gene define a subgroup with a better outcome, with their presence able to counterbalance the deleterious effect of the *FLT3*-ITD mutations, when coexpressed.<sup>7</sup> The interaction between the *NPM1* and *FLT3*-ITD allelic ratio discriminates groups with different prognosis.<sup>2</sup> Whereas the *NPM1* mutation (*NPM1*<sup>mut</sup>) alone has retained its favorable prognostic role, the genotypes resulting from the combination of *NPM1*<sup>mut</sup> and *FLT3*-ITD with a “low” (*FLT3*<sup>low</sup>) or “high” (*FLT3*<sup>high</sup>) allelic ratio are now distributed in all three ELN2017 risk categories.<sup>8,9</sup> Furthermore, the 2017 version of the ELN recommendations was further updated in that they included in the adverse group several new gene mutations such as *RUNX1*, *ASXL1*, and *TP53*.<sup>2</sup>

Recently, the Gruppo Italiano Malattie EMatologiche dell'Adulto (GIMEMA) Foundation accomplished a prospective, multicenter clinical trial (GIMEMA AML1310). The study strategy relied on the prognostic integration of pretreatment cytogenetics and genetics with postconsolidation measurable residual disease (MRD) as detected by using multiparametric flow cytometry (MFC). Based on this strategy, patients were to receive postconsolidation autologous stem cell transplant (auto-SCT) or allogeneic stem cell transplant (allo-SCT), respectively, depending on their risk profile.<sup>9-12</sup> At the time of study design, the National Comprehensive Cancer Network (NCCN) 2009 risk classification<sup>13</sup> was the only one available, and patients were therefore stratified according to this guideline. Accordingly, all patients carrying an *FLT3* mutation were considered as high risk and allocated to allo-SCT, regardless of other biological features, including *NPM1* mutation status, *FLT3* allelic ratio, or any concurrent genetic/cytogenetic abnormality. Furthermore, when the study was activated, no *FLT3* inhibitors had yet been licensed, nor were they administered to patients enrolled in the trial.<sup>14</sup>

Five hundred patients were included in the final analysis of the AML1310 trial, and the combination of baseline genetics with the assessment of the allelic ratio in *FLT3*-ITD-positive (*FLT3*-ITD<sup>mut</sup>) cases allowed us to reclassify 445 of these 500 patients according to the new ELN2017 risk classification. At the writing of the current analysis, the mutational sequencing of *RUNX1*, *ASXL1*, and *TP53* genes by next-generation sequencing is not yet available.<sup>15</sup>

The aim of the current post hoc analysis was to validate the ELN2017 classification in a prospective series of homogeneously treated patients.

## Patients and methods

### Patients

Previously untreated patients with a diagnosis of de novo AML according to the World Health Organization diagnostic criteria<sup>16</sup> were eligible for the GIMEMA AML1310 Study (EudraCT number 2010-023809-36; ClinicalTrials.gov identifier #NCT01452646). Main inclusion and exclusion criteria have been published elsewhere<sup>9</sup> and are detailed in the supplemental Materials. All participants gave their informed consent, and the study was conducted in accordance with the Declaration of Helsinki after approval by the ethics committees of the participating hospitals/academic institutions.

### Study design

The main objective of the AML1310 study was to verify whether the delivery of a postremission therapy, the intensity of which was risk driven, extended the 2-year OS compared with the historic data from the previous LAM99P GIMEMA trial.<sup>17</sup> Upfront evaluation included bone marrow sampling for morphology, cytogenetics, molecular genetics, and MFC analysis. The baseline MFC assessment was a necessary step to identify leukemia-associated immunophenotypes. Identification of baseline leukemia-associated immunophenotypes by using a high-sensitivity 8-color MFC assay was the essential requirement for monitoring MRD after therapy at the established postconsolidation time point. Based on several retrospective validations in the context of former European Organisation for Research and Treatment of Cancer (EORTC)/GIMEMA protocols,<sup>12</sup> the threshold for discriminating MRD-negative cases from MRD-positive cases was set at  $3.5 \times 10^{-4}$  (0.035%) residual leukemic cells, upon full blood count recovery. As a mandatory step, patients were studied, at diagnosis, for the presence of mutations of *NPM1*, *FLT3*-ITD, *FLT3*-TKD, and *c-KIT* and of rearrangements *RUNX1*-*RUNX1T1* or *CBFβ*/*MYH11*, defining core-binding factor (*CBF*)-positive AML. Using polymerase chain reaction, the threshold discriminating low from high *FLT3*-ITD allelic ratio was set at 0.5, as proposed in the ELN2017 recommendations.<sup>2,18</sup> Baseline cytogenetic, *CBF* rearrangements, *NPM1*, *FLT3* mutation status, and the allelic ratio for *FLT3*-ITD-positive cases were considered necessary information to classify each patient according to the ELN2017 risk stratification.<sup>7,19,20</sup>

All molecular and MFC analyses were centralized at Laboratorio di Diagnostica Integrata Oncoematologica (Tor Vergata University Hospital, Rome, Italy), whereas baseline conventional karyotyping was conducted at local institutions. Bone marrow and peripheral blood were used as sources to assess response to treatment, according to the recommendations of an international working group.<sup>21</sup> The AML1310 trial was designed at a time when ELN 2010/2017 recommendations were not available. Therefore, when the trial regulatory path was concluded, patients were recruited and stratified according to the contemporary NCCN2009 version 1 classification.<sup>13</sup> Induction and consolidation regimens have been reported elsewhere<sup>9</sup> and in the supplemental Materials. By integrating the NCCN 2009 classification with the level of postconsolidation MRD, 4 categories of risk were considered: favorable-risk (NCCN-FR) or poor-risk (NCCN-PR) patients, who were submitted to auto-SCT or allo-SCT, respectively; intermediate MRD-negative or intermediate MRD-positive patients, who were to receive auto-SCT or allo-SCT. allo-SCT and auto-SCT were to be performed within 3 months of the end of the consolidation course.

## Statistical analysis and sample size calculation

The primary study objective was the percentage of OS at 2 years. An estimated number of 213 subjects were initially required to accomplish this primary objective. This sample size was to achieve a 90% power to detect a difference of 10% between the null hypothesis that OS at 2 years is 50% and the alternative hypothesis that OS is 60%, using a single-stage phase 2 design with a 5% significance level (based on data of the historic control group GIMEMA LAM99P).<sup>17</sup> Based on the historic control group, we also considered that ~70% of the observed patients would have been classified as intermediate risk, thereby allowing us to reach the figure of 150 patients available for MRD-driven treatment allocation. However, after 173 subjects were enrolled, only 56 belonged to the intermediate-risk category (32% vs 70% expected). Thus, to reach the target of 150 subjects belonging to the intermediate-risk category, an amendment to the protocol was adopted in 2013, and the sample size was adjusted to 515 subjects to recruit.

Patient and disease characteristics were summarized by means of cross-tabulations for categorical variables or by quintiles for continuous variables. In univariate analysis, nonparametric tests were performed for comparisons among the ELN2017 and combined NPM1/FLT3-ITD groups ( $\chi^2$  test and Fisher's exact test in case of categorical variables or response rate, Mann-Whitney and Kruskal-Wallis tests in case of continuous variables).

OS (time elapsed from treatment start to death) and disease-free survival (DFS) (time from complete remission [CR] to relapse or death in remission) were calculated by using the Kaplan-Meier product limit estimator. Differences in terms of OS and DFS were evaluated by means of log-rank tests in univariate analysis, and by means of a Cox regression model in multivariate analysis, after assessment of proportionality of hazards. All variables with a *P* value < .15 in univariate analysis were considered in the multivariate models. The influence of the transplant on the survival outcome was evaluated in the Cox model by means of a time-dependent covariate. Confidence intervals (CIs) were calculated at the 95% level, and all tests were two-sided, accepting *P* ≤ .05 as indicating a statistically significant difference. All analyses were performed by using SAS version 9.4 (SAS Institute, Inc., Cary, NC) and R (R Foundation for Statistical Computing, Vienna, Austria) system software. Study data were collected and managed by using the REDCap<sup>20</sup> electronic data capture tools hosted at the GIMEMA Foundation.

## Results

Overall, 515 patients from 55 GIMEMA institutions were registered to the AML1310 trial. Fifteen patients were considered ineligible because of infections and/or death before treatment initiation, and thus the final analysis included 500 individuals. Median age was 49 (18-61) years, with a slight male predominance (52%). Table 1 summarizes patient characteristics.

Among 500 cases, *FLT3-ITD* and *NPM1* mutations were detected in 123 (25.8%) and in 182 (38%) cases, respectively. Of 123 *FLT3-ITD*<sup>mut</sup> and 182 *NPM1*<sup>mut</sup>, coexpression was observed in 80 patients. The *FLT3-ITD* allelic ratio, available in 109 (88.6%) of 123 patients, was low and high in 40 (32.5%) and 69 (56.1%) patients, respectively. Patients arranged per the original, per-protocol risk stratification were as follows: 138 (27.6%) were NCCN-FR, 174 (34.8%) were NCCN-IR, and 188 (37.6%) were NCCN-PR. The

**Table 1. General characteristic of the study population**

Characteristic	Value
No.	500
Age, median (range), y	49 (18-61)
<b>Sex, n (%)</b>	
Male	260 (52)
Female	240 (48)
<b>Cytogenetics, n (%)</b>	
FR	48 (11)
IR	316 (73)
AR	68 (16)
RUNX1/RUNX1T1, n (%)	27 (5)
CBFB/MYH1, n (%)	37 (7)
<b>FLT3-ITD<sup>mut</sup>, n (%)</b>	
FLT3 <sup>low</sup>	69 (14)
FLT3 <sup>high</sup>	40 (8)
<b>NPM1<sup>mut</sup>, n (%)</b>	182 (38)
NPM1 <sup>mut</sup> /FLT3-ITD <sup>mut</sup> , n (%)	80 (16)
<b>NCCN, n (%)</b>	
NCCN-FR	138 (28)
NCCN-IR	174 (34)
NCCN-PR	188 (38)
<b>ELN2017, n (%)</b>	
ELN2017-FR	186 (37)
ELN2017-IR	179 (36)
ELN2017-AR	80 (16)
ELN2017-NC	55 (11)

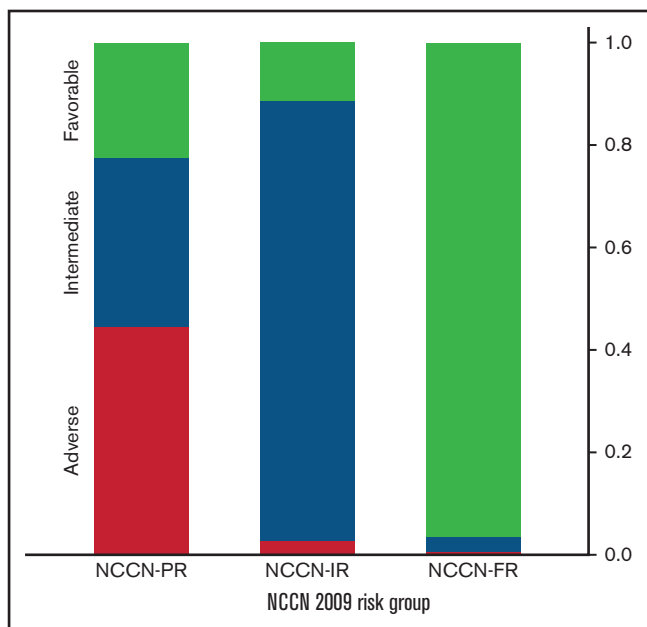
conversion from NCCN2009 to ELN2017 risk stratification resulted in the loss of 55 patients in whom ELN2017 allocation was not feasible due to lack of cytogenetics or *FLT3-ITD* allelic ratio assessment. These 55 patients were grouped in a nonclassifiable category named ELN2017-NC. Therefore, 445 (89%) of 500 patients were ELN2017 reclassified as follows: 186 (41.8%) were favorable risk (ELN2017-FR), 179 (40.2%) intermediate risk (ELN2017-IR), and 80 (18%) adverse risk (ELN2017-AR). By adding up estimation of the allelic ratio, *FLT3-ITD*-positive patients were spread across all three ELN2017 categories: 38 (31%) in the ELN2017-FR group, 51 (41%) in the ELN2017-IR group, and 20 (16%) in the ELN2017-AR group. Four additional patients were moved from the NCCN2009-PR category to the ELN2017-IR category due to the reclassification of the cytogenetic abnormality. Patients' risk categorization according to ELN2017 evidenced a variable redistribution of NCCN2009-PR cases across all ELN2017 risk groups, with 38 (20.2%), 55 (29.2%), and 20 (10.6%) of 188 NCCN2009-PR patients now being reclassified as ELN2017-FR, ELN2017-IR, and ELN2017-NC. Based on this outcome, the analysis of the differential distribution of patients resulted in a 65.4% concordance (327 of 500) between the 2 risk stratification models. Indeed, only 75 (39%) of 188 cases from the NCCN-PR group were categorized as adverse risk also according to ELN2017. At variance, a higher degree of concordance was observed between the NCCN-FR/ELN2017-FR and NCCN-IR/ELN2017-IR groups, with 132 (95.6%) of 138 NCCN-FR cases and 120 (68.9%) of 174 NCCN-

IR cases still being classified as ELN2017-FR and ELN2017-IR (Figure 1).

Overall, after a maximum of 2 cycles of induction, 361 (72%) patients obtained a CR/CR with incomplete blood count recovery: 163 (88.1%), 114 (65%), 45 (56.2%), and 39 (70%) in the ELN2017-FR, ELN2017-IR, ELN2017-AR, and ELN2017-NC groups, respectively ( $P < .001$ ). The majority of these patients (342 of 361 [95%]) successfully started the consolidation phase and were subsequently allocated to the corresponding transplant procedure: 177 (52%) to auto-SCT and 165 (48%) to allo-SCT. Of the 177 auto-SCT candidates, 111 (62.7%) underwent transplant (82 [73.9%] ELN2017-FR, 19 [17.1%] ELN2017-IR, 2 [1.8%] ELN2017-AR, and 8 [7.2%] ELN2017-NC). The number of allo-SCT candidates, with the addition of 23 patients who achieved a CR after salvage therapy, was further increased to 188. Overall, of the 188 allo-SCT candidates, 132 (70.2%) underwent transplant (25 [18.9%] ELN2017-FR, 61 [46.2%] ELN2017-IR, 25 [18.9%] ELN2017-AR, and 21 [15.9%] ELN2017-NC) ( $P < .001$ ). Finally, 19 patients who did not undergo auto-SCT or allo-SCT received additional high-dose cytarabine chemotherapy: 18 from the ELN2017-FR group and 1 from the ELN2017-IR group.

### Survival analyses according to ELN2017

After a median follow-up of 28.8 months, 2-year OS and DFS were 56% (95% CI, 52-61; median duration, 38 months) and 54% (95% CI, 49-60; median duration, 32.4 months), respectively.

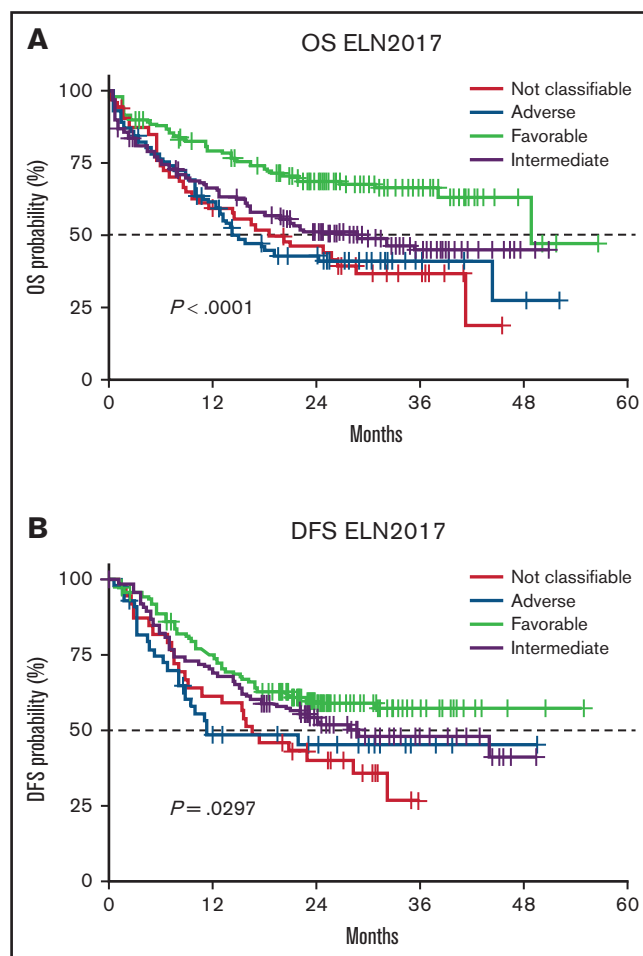


**Figure 1.** Spine plot for each NCCN showing the proportion of adverse (red), intermediate (blue), and favorable (light green) risk. A small proportion of NCCN2009 patients (75 of 188 [39%]) remained classified as high risk according to ELN2017. A high proportion of NCCN2009-PR cases were redistributed across all ELN2017 risk groups, with 38 (20.2%), 55 (29.2%), and 20 (10.6%) of 188 NCCN2009-PR patients now being reclassified as ELN2017-FR, ELN2017-IR, and ELN2017-NC, respectively. At variance, 132 (95.6%) of 138 NCCN-FR cases and 120 (68.9%) of 174 NCCN-IR cases remained classified as ELN2017-FR and ELN2017-IR, respectively.

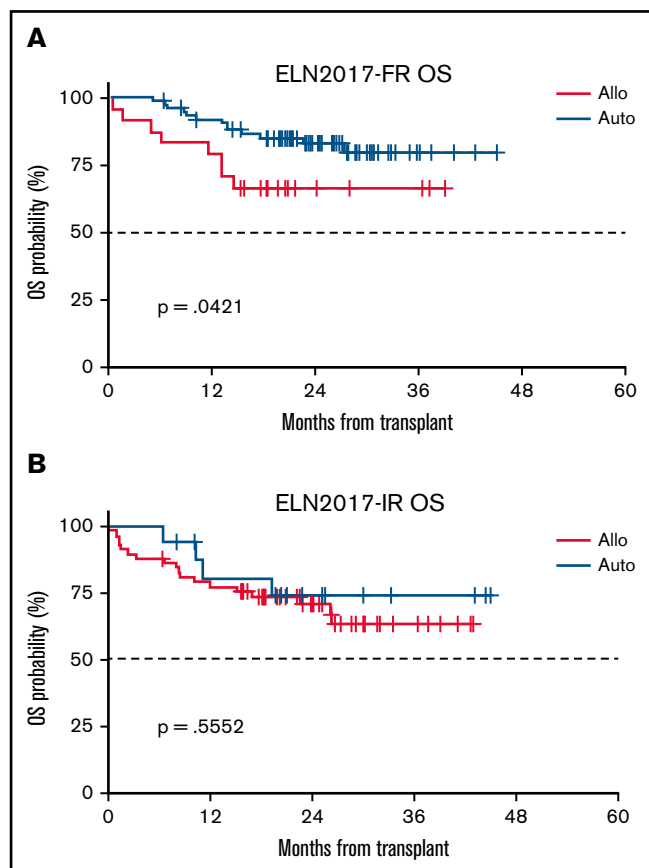
Cumulative incidence of relapse was 33% (95% CI, 28-38) (supplemental Figure 1).

Stratification of the study population according to ELN2017 risk classification gave rise to 3 groups with a significantly different duration of OS: ELN2017-FR patients, who had the best outcome (2-year OS, 68.8%), followed by ELN2017-IR (2-year OS, 51.3%) and ELN2017-AR (2-year OS, 42.8%) patients. ELN2017-NC patients had an intermediate outcome, with their 2-year OS of 45.8% ( $P < .001$ ) (Figure 2A). Two-year DFS was 59.9%, 54.2%, 45.5%, and 40.3% for the ELN2017-FR, ELN2017-IR, ELN2017-AR, and ELN2017-NC patients, respectively ( $P = .0297$ ) (Figure 2B). Two-year cumulative incidence of relapse was 31.3%, 29.4%, 42.8%, and 39.2% for the ELN2017-FR, ELN2017-IR, ELN2017-AR, and ELN2017-NC patients ( $P = .2343$ ) (supplemental Figure 2).

We next investigated the impact of the selected postconsolidation strategy (allo-SCT vs auto-SCT) on each ELN2017 category. Due to the AML1310 protocol design and the consequent negligible number of patients receiving auto-SCT ( $n = 2$ ) in the ELN2017-AR



**Figure 2.** Patient outcome according to ELN2017 risk stratification and FLT3/NPM1 gene interactions. (A) Two-year OS was 68.8%, 51.3%, 42.8%, and 45.8% for patients belonging to the ELN2017-FR, ELN2017-IR, ELN2017-AR, and ELN2017-NC categories, respectively. (B) Two-year DFS was 59.9%, 54.2%, 45.5%, and 40.3% for the ELN2017-FR, ELN2017-IR, ELN2017-AR, and ELN2017-NC patients.



**Figure 3. Correlation between postconsolidation strategy and outcome for each ELN2017 risk category.** (A) Benefit of auto-SCT in the ELN2017-FR category (2-year OS of 83.3% vs 66.7% for auto-SCT vs allo-SCT). (B) Almost equal performance of allo-SCT and auto-SCT in the ELN2017-IR category (2-year OS of 70.8% and 73.9%, for allo-SCT and auto-SCT, respectively).

group, comparison between the 2 transplant procedures was restricted to the ELN2017-FR and ELN2017-IR categories. In the ELN2017-IR category, we did not observe any difference in terms of OS duration between patients who received allo-SCT and those who received auto-SCT (2-year OS of 70.8% and 73.9%, for allo-SCT and auto-SCT, respectively;  $P = .5552$ ). However, categorized as ELN2017-FR and being submitted to auto-SCT rather than to allo-SCT were associated with a survival advantage (2-year OS of 83.3% vs 66.7% for auto-SCT and allo-SCT;  $P = .0421$ ) (Figure 3). Among the ELN2017-IR category, nonrelapse mortality (NRM) for patients submitted to auto-SCT or allo-SCT was 5.3% and

13.7% ( $P = .3246$ ). In the ELN2017-FR category, NRM was 0% and 16.7% ( $P = .0002$ ) for those receiving auto-SCT or allo-SCT.

Considering that, in the original AML1310 protocol, the postconsolidation treatment of the NCCN2009-IR category was decided according to MRD status, we analyzed if MRD maintains its role in the newly designed ELN2017-IR category. Postconsolidation MRD data were available in 85 patients, 55 negative (64.7%) and 30 positive (36.3%). In the overall population, MRD-negative patients had a better 2-year OS (76.5% vs 58.8%), although it was not significant ( $P = .247$ ). Among MRD-negative patients, 26 underwent allo-SCT and 10 auto-SCT, respectively; among MRD-positive patients, 17 underwent allo-SCT and 3 auto-SCT. The 2-year OS, when landmarked “from” transplant, did not differ between auto-SCT and allo-SCT in MRD-negative patients (85.7% vs 77.8%;  $P = .234$ ); among MRD-positive patients, it was significantly longer for those receiving allo-SCT (75% vs 0%;  $P = .0231$ ) (supplemental Figure 3).

Univariate analyses provided evidence regarding the independent role of each considered covariate in influencing duration of OS. Indeed, compared with the ELN2017-FR group, belonging to the ELN2017-AR (hazard ratio, 2.203; CI, 1.496-3.246;  $P < .0001$ ), ELN2017-IR (hazard ratio, 1.796; CI, 1.293-2.494;  $P = .0005$ ), and ELN2017-NC (hazard ratio, 2.267; CI, 1.488-3.455;  $P = .0001$ ) category was associated with a shorter duration of OS.

The multivariate model for OS prediction confirmed the significant prognostic role of age (supplemental Figure 4), ELN2017 risk stratification, and transplantation as time-dependent parameters (Table 2).

## Discussion

In the current analysis, we showed that, when implemented in the prospective, risk-adapted, MRD-driven, AML1310 GIMEMA protocol, ELN2017 risk classification maintained its prognostic significance, thereby affecting OS and DFS.

Prognostic stratification of AML has been substantially improved by the introduction of the ELN2017 recommendations. Indeed, through the combination of a broad range of genetic and cytogenetic abnormalities, the ELN2017 recommendations identify 3 prognostic classes (favorable, intermediate, and adverse). The ELN2017 version, compared with the ELN2010 precedent edition, is influenced by evidence that the mutual interactions between *NPM1* and *FLT3*, and the *FLT3-ITD* allelic ratio, play a crucial role in discriminating patients with different outcomes. Based on this, the 2017 revisited edition has resulted in a redistribution of *FLT3-ITD* mutations across all 3 risk categories: cases of *NPM1*<sup>mut</sup>/*FLT3-ITD*<sup>low</sup> have been

**Table 2. Multivariate model for OS prediction**

Parameter	Detail	Probability ( $\chi^2$ )	Hazard ratio	95% Lower CI	95% Upper CI
Age		<.0001	1.033	1.019	1.048
ELN2017 risk group	ELN2017-NC vs ELN2017-FR	.0003	2.187	1.434	3.334
ELN2017 risk group	ELN2017-AR vs ELN2017-FR	<.0001	2.187	1.481	3.229
ELN2017 risk group	ELN2017-IR vs ELN2017-FR	.0003	1.838	1.321	2.557
Transplant, covariate time-dependent	Transplant vs no transplant	.0185	0.674	0.485	0.936

allocated in the favorable-risk group, those of *NPM1*<sup>mut</sup>/*FLT3-ITD*<sup>high</sup> or *NPM1*<sup>wt</sup>/*FLT3-ITD*<sup>low</sup> in the intermediate-risk group, and finally cases of *NPM1*<sup>wt</sup>/*FLT3-ITD*<sup>high</sup> in the adverse-risk group. The reliability of this risk stratification has been widely recognized, leading to its incorporation into the NCCN clinical practice guidelines for AML.<sup>22</sup>

Patients recruited to the GIMEMA AML1310 trial were categorized based on the NCCN2009 criteria.<sup>13</sup> Therefore, the purpose of this post hoc analysis was to verify if the prospective, risk-adapted strategy of GIMEMA AML1310 trial held true when the same population was retrospectively rearranged according to the ELN2017 classification.

The exercise of a retrospective application of ELN2017 criteria to the AML1310 cohort was feasible in ~90% of the patients. As a consequence of the new role assigned to the *FLT3-ITD* allelic ratio and to the concomitant expression of *NPM1* in the wild-type or mutated conformation, many patients originally allocated into the NCCN-PR risk category were reassigned across all three ELN2017 subgroups.

Once the small group of 55 ELN2017-NC patients (11% of the whole cohort) were excluded, an overall concordance of 65.4% was observed between the ELN2017 and the original per-protocol NCCN2009 categorization. The overlap was more evident for the NCCN-FR/ELN2017-FR and NCCN-IR/ELN2017-IR groups, with less concordance between NCCN-PR/ELN2017-AR cases. When we focused on CR rate, ELN2017 offered a more consistent picture of the risk profile of the 3 categories. Indeed, while for NCCN-FR/ELN2017-FR and NCCN-IR/ELN2017-IR patients the CR rate was equivalent (88.0% and 65.0%, respectively), those reassigned to the ELN2017-AR category had a CR rate of 56.2%, compared with 69.9% for NCCN-PR. The CR rate achieved by the category of ELN2017-AR patients seems more logically connected to the risk profile of this group than the CR frequency shown for the NCCN2009-PR group. This observation seems even more robust if one considers the prospective nature of the AML1310 trial and then the homogeneous induction and consolidation chemotherapy that was delivered.

At variance with the original NCCN2009-driven protocol design, the re-allocation of the cases according to the ELN2017 classification resulted in an ELN2017-FR group characterized by a mix of patients who received allo-SCT or auto-SCT. Similarly, the ELN2017-IR category also included a composite population of patients because, as per protocol, the transplant option was decided according to the level of MRD after consolidation. By doing so, robust cost/benefit implications came across that were not so quickly captured in origin when the NCCN2009 classification was used. In fact, in the ELN2017-FR category, auto-SCT was associated with a better 2-year OS compared with that of allo-SCT. We assume that the favorable outcome of patients within the ELN2017-FR category was jeopardized by the NRM of allo-SCT. Conversely, in the ELN2017-IR category, auto-SCT and allo-SCT were associated with an equivalent duration of OS. In fact, among the 120 patients classified as intermediate risk both according to the ELN2017 and the NCCN2009 risk classifications, the allo-SCT counterbalances the poor prognosis of an MRD-positive status and prolongs the OS of MRD-positive patients to equalize the one of MRD-negative patients. This confirms that postconsolidation MRD assessment, regardless

of the technique used, remains critical in the intermediate risk category to inform the decision-making process of transplant allocation.

Risk stratification according to ELN2017 was not feasible in 55 cases, accounting for ~11% of the whole series. Notwithstanding, we kept this ELN2017-NC group of patients in the analysis for comparison purposes. Patients belonging to this category had a CR rate similar to the ELN2017-IR one but an OS and DFS similar to the ELN2017-AR one. Because almost one-half of these patients (26 of 55 [47%]) were not submitted to either auto-SCT or allo-SCT, we speculate that their dismal outcome could be attributed to the delivery of a suboptimal postconsolidation therapy.

We are aware of the limitations of our study, mainly attributable to the post hoc nature of the analysis. The GIMEMA AML1310 trial was not designed for the purpose of the current analysis, and at the time of study design not even the 2010 version of the ELN recommendations was available. Furthermore, we lack information about the mutational status of genes such as *RUNX1*, *ASXL1*, and *TP53*, and this may have led to a misclassification of some patients of our cohort. However, an extensive next-generation sequencing analysis of the present series is ongoing. Unfortunately, the results are not yet available.

In conclusion, our attempt to reclassify, according to the ELN2017 criteria, a large cohort of patients with AML, who were originally risk stratified according the NCCN2009 classification, resulted in a more consistent prediction of outcome. In the context of the GIMEMA AML1310 trial, these patients received homogeneous induction and consolidation courses, with the postconsolidation program being decided on baseline genetics/cytogenetics and post-consolidation MRD assessment. In this context, the ELN2017 classification was able to segregate, even better than NCCN2009, cohorts of patients with distinct clinical outcome. This is due to the impact of the *FLT3-ITD* allelic ratio and interaction between *NPM1* and *FLT3-ITD* in refining the prognostic value of the *FLT3-ITD* category.

We believe that the current analysis represents an effective example of integration of a modern risk-adapted therapeutic program and a modern risk-stratified approach such as ELN2017. The expanding knowledge of new gene mutations and the ever more diffuse availability of targeted agents will contribute to enhancing the outcome prediction of the ELN2017 classification.

## Acknowledgment

This work was supported by AIRC 5 × 1000 call “Metastatic disease: the key unmet need in oncology” to MYNERVA project, #21267 (MYeloid NEoplasms Research Venture AIRC. A detailed description of the MYNERVA project is available at <https://progettomynerva.it/>).

## Authorship

Contribution: F.B., R.P., A.V., A.P., and V.A. created the concept and design; A.C., L. Melillo, V. C., R.C., P.D.F., G.S., P.S., F.L., G.M., M.L., S.C., L. Maurillo, M.I.D.P., G.P., M.A.I.-C., T.O., and S.L. provided the study materials or patients; F.B., R.P., A.P., and V.A. collected and assembled the data; and F.B., L. Maurillo, M.I.D.P., M.A.I.-C., T.O., and S.L. performed the laboratory testing and monitoring; and all authors contributed to the writing and the final approval of the manuscript.

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

ORCID profiles: F.B., 0000-0003-4320-9253; R.P., 0000-0003-2855-8121; A.P., 0000-0001-8648-885X; V.A., 0000-0003-2782-2752; P.D.F., 0000-0002-1835-0581; L.M., 0000-0001-7297-

4988; G.P., 0000-0002-8944-0795; M.T.V., 0000-0002-6164-4761; A.V., 0000-0002-0245-0553.

Correspondence: Francesco Buccisano, Department of Biomedicine and Prevention, Tor Vergata University of Rome, Via Montpellier 1, Rome, Italy; e-mail: francesco.buccisano@uniroma2.it.

## References

1. Döhner H, Estey EHE, Amadori S, et al; European LeukemiaNet. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood*. 2010;115(3):453-474.
2. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424-447.
3. Patel JP, Gönen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N Engl J Med*. 2012;366(12):1079-1089.
4. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med*. 2016;374(23):2209-2221.
5. Schlenk RF, Kayser S, Bullinger L, et al; German-Austrian AML Study Group. Differential impact of allelic ratio and insertion site in FLT3-ITD-positive AML with respect to allogeneic transplantation. *Blood*. 2014;124(23):3441-3449.
6. Small D. FLT3 mutations: biology and treatment. *Hematology Am Soc Hematol Educ Program*. 2006;2006:178-184.
7. Ivey A, Hills RK, Simpson MA, et al; UK National Cancer Research Institute AML Working Group. Assessment of minimal residual disease in standard-risk AML. *N Engl J Med*. 2016;374(5):422-433.
8. Gale RE, Green C, Allen C, et al; Medical Research Council Adult Leukaemia Working Party. The impact of FLT3 internal tandem duplication mutant level, number, size, and interaction with NPM1 mutations in a large cohort of young adult patients with acute myeloid leukemia. *Blood*. 2008;111(5):2776-2784.
9. Venditti A, Piciocchi A, Candoni A, et al. GIMEMA AML1310 trial of risk-adapted, MRD-directed therapy for young adults with newly diagnosed acute myeloid leukemia. *Blood*. 2019;134(12):935-945.
10. Buccisano F, Maurillo L, Spagnoli A, et al. Cytogenetic and molecular diagnostic characterization combined to postconsolidation minimal residual disease assessment by flow cytometry improves risk stratification in adult acute myeloid leukemia. *Blood*. 2010;116(13):2295-2303.
11. Buccisano F, Maurillo L, Del Principe MI, et al. Minimal residual disease as a biomarker for outcome prediction and therapy optimization in acute myeloid leukemia. *Expert Rev Hematol*. 2018;11(4):307-313.
12. Maurillo L, Buccisano F, Del Principe MI, et al. Toward optimization of postremission therapy for residual disease-positive patients with acute myeloid leukemia. *J Clin Oncol*. 2008;26(30):4944-4951.
13. National Comprehensive Cancer Network. Acute myeloid leukemia. Version 1. NCCN Clinical Practice Guidelines in Oncology. 2009. [https://www.nccn.org/professionals/physician\\_gls/pdf/aml.pdf](https://www.nccn.org/professionals/physician_gls/pdf/aml.pdf). Accessed 6 May 2018.
14. Stone RM, Mandrekar SJ, Sanford BL, et al. Midostaurin plus chemotherapy for acute myeloid leukemia with a FLT3 mutation. *N Engl J Med*. 2017;377(5):454-464.
15. Döhner K, Thiede C, Jahn N, et al. Impact of NPM1/FLT3-ITD genotypes defined by the 2017 European LeukemiaNet in patients with acute myeloid leukemia. *Blood*. 2020;135(5):371-380.
16. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009;114(5):937-951.
17. Lo-Coco F, Cuneo A, Pane F, et al; Acute Leukemia Working Party of the GIMEMA Group. Prognostic impact of genetic characterization in the GIMEMA LAM99P multicenter study for newly diagnosed acute myeloid leukemia. *Haematologica*. 2008;93(7):1017-1024.
18. Thiede C, Studel C, Mohr B, et al. Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis. *Blood*. 2002;99(12):4326-4335.
19. Gorello P, Cazzaniga G, Alberti F, et al. Quantitative assessment of minimal residual disease in acute myeloid leukemia carrying nucleophosmin (NPM1) gene mutations. *Leukemia*. 2006;20(6):1103-1108.
20. Gabert J, Beillard E, van der Velden VHJ, et al. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia – a Europe Against Cancer program. *Leukemia*. 2003;17(12):2318-2357.
21. Cheson BD, Bennett JM, Kopecky KJ, et al; International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia [published correction appears in *J Clin Oncol*. 2004;22(3):576]. *J Clin Oncol*. 2003;21(24):4642-4649.
22. Tallman MS, Wang ES, Altman JK, et al; OCN. Acute Myeloid Leukemia, Version 3.2019, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2019;17(6):721-749.