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

Tracking no: ADV-2021-005717R2

Francesco Buccisano (Tor Vergata University of Rome, Italy) Raffaele Palmieri (Ematologia, Dipartimento di Biomedicina e Prevenzione, Università di Roma "Tor Vergata", Italy) Alfonso Piciocchi (GIMEMA Foundation, Italy) Valentina Arena (GIMEMA Foundation, Rome, Italy) Anna Candoni (Division of Hematology, University of Udine, Italy) Lorella Melillo (Fondazione IRCCS Casa Sollievo della Sofferenza, Udine, Italy) Valeria Calafio (AOU Policlinico-Vittorio Emanuele - Department of Hematology, Italy) Roberto Cairoli (ASST Grande Ospedale Metropolitano Niguarda, Italy) Paolo De Fabritiis (Sant' Eugenio Hospital, Italy) Gabriella Storti (AORN ' San G. Moscati ' Avellino, Italy) Prassede Salutari (ospedale civile pescara, Italy) Francesco Lanza (HEMATOLOGY INSTITUTE, Italy) Giovanni Martinelli (Policlinico S. Orsola-Malpighi, Bologna, Italy) Mario Luppi (University of Modena and Reggio Emilia, Italy) Saveria Capria (Policlinico Umberto I, Sapienza University, Italy) Luca Maurillo (Fondazione Policlinico Tor Vergata, Italy) Maria Ilaria Del Principe (FONDAZIONE POLICLINICO TOR VERGATA, Italy) Giovanniacinto Paterno (Università di Tor Vergata, Roma, Italy) Maria Antonietta Irno Consalvo (Università di Tor Vergata, Italy) Tiziana Ottone (Università di Tor Vergata, Roma, Italy) Serena Lavorgna (Università Tor Vergata Roma, Italy) Maria Teresa Voso (Universita' di Roma Tor Vergata, Italy) Paola Fazi (GIMEMA Foundation, Italy) Marco Vignetti (Dept of Cellular Biotechnologies and Hematology, University "La Sapienza", Italy) William Arcese (Rome Transplant Network, Italy) Adriano Venditti (Fondazione Policlinico Tor Vergata, Italy)

Abstract:
The 2017 version of the ELN recommendations, by integrating cytogenetics and mutational status of specific genes, sort out 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 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 autologous stem cell transplant (AuSCT) if categorized as favorable-risk or autologeneic stem cell transplant (ASCT) if adverse-risk. Intermediate-risk pts were to receive AuSCT or ASCT based on the post-consolidation levels of Measurable Residual Disease as measured by flow-cytometry. Risk categorization was originally conducted according to NCCN2009 recommendations. Among 500 patients, 445 (89%) were re-classified according to the ELN2017 criteria: ELN2017-FR (186/455; 41.8%), ELN2017-IR (179/445 40.2%) and ELN2017-AR (80/455; 18%); in 55 patients (11%) ELN2017 was not applicable (ELN2017-NC). Two-year overall survival (OS) was 68.8%, 51.3%, 45.8% and 42.8% for ELN2017-FR, ELN2017-IR, ELN2017-NC, and ELN2017-AR group, respectively (p<0.001). When comparing the two different transplant strategies in each ELN2017 risk category, a significant benefit of AuSCT over ASCT was observed among ELN2017-FR patients (2-years OS of 83.3% vs. 66.7%; p=0.0421). The two transplant procedures performed almost equally in the ELN2017-IR group (2-years OS of 73.9% vs. 70.8%; p=0.5552). 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.

Conflict of interest: No COI declared

COI notes:

Preprint server: No;

Author contributions and disclosures: F.B., R.P., A.V., A.P., and V.A., created the concept and design; A.C., L.Me, V.C., R.C., P. D.F., G.S., P.S., F.L., G.M., M.L., S.C., L.Ma, 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; F.B., L.Ma, 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 of the manuscript and to the final approval of the manuscript. F.B. and R.P. equally contributed to the manuscript

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: For original AML1310 trial
Clinical trial registration information (if any):
ELN2017 risk stratification improves outcome prediction when applied to the prospective, GIMEMA AML1310 protocol.

ELN 2017 risk to guide treatment strategy in AML

Francesco Buccisano1*, Raffaele Palmieri1*, Alfonso Piciocchi2, Valentina Arena2, Anna Candoni3, Lorella Melillo4, Valeria Calafiore5, Roberto Cairoli6, Paolo de Fabritiis7, Gabriella Storti8, Prassede Salutari9, Francesco Lanza10, Giovanni Martinelli11,12, Mario Luppi13, Saveria Capria14, Luca Maurillo1, Maria Ilaria Del Principe1, Giovangiacinto Paterno1, Maria Antonietta Irno-Consalvo1, Tiziana Ottone1, Serena Lavorgna3, Maria Teresa Voso1, Paola Fazi2, Marco Vignetti2, William Arcese1,15, Adriano Venditti1.

1 Ematologia, Dipartimento di Biomedicina e Prevenzione, Università di Roma “Tor Vergata”, Roma,
2 GIMEMA Foundation, Rome, Italy
3 Hematology, Azienda Sanitaria Universitaria Integrata di Udine, Udine, Italy
4 Fondazione IRCCS Casa Sollievo della Sofferenza, UO di Ematologia, San Giovanni Rotondo (FG), Italy
5 Ospedale Ferrarotto, Catania, Italy
6 Milano: Roberto Cairoli; Ospedale Niguarda Ca Granda, Milan, Italy
7 S. Eugenio: Paolo de Fabritiis; Ospedale S. Eugenio, Rome, Italy
8 Avellino: Gabriella Storti; Azienda Ospedaliera S.G. Moscati, Avellino, Italy
9 Pescara: Prassede Salutari; Azienda USL di Pescara, Pescara, Italy
10 Ravenna: Francesco Lanza; Ospedale S. Maria delle Croci, Ravenna, Italy
11 Istituto Tumori della Romagna, Meldola, Italy
12 Policlinico S. Orsola-Malpighi, Bologna, Italy
13 Mario Luppi; Ematologia, Dipartimento di Scienza Mediche e Chirurgiche Materno-Infantili e dell’Adulto, Università degli studi di Modena e Reggio Emilia, Modena, Italy
14 Roma Sapienza: Saveria Capria; Dipartimento di Biotecnologie Cellulari ed Ematologia, Ematologia Università degli Studi Sapienza, Rome, Italy
15 Rome Transplant Network, Roma, Italy

* The authors equally contributed to the manuscript

KEY POINTS:
1. Within homogeneous ELN risk categories, different post-remission approaches (AuSCT vs ASCT) result in different outcomes
2. MRD confirms its role as a driver of transplant allocation for patients placed in the ELN2017 intermediate risk category.

Correspondence: Francesco Buccisano, MD; e-mail: francesco.buccisano@uniroma2.it; Tel.: +39-06-20903228; Fax.: +39-06-20903212.
Acknowledgements: This work was supported by AIRC 5x1000 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/).

Abstract: 249 words

Manuscript: 3402 words

Figures: 3

Tables: 2

References: 22
ABSTRACT

The 2017 version of the ELN recommendations, by integrating cytogenetics and mutational status of specific genes, sort out 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 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 autologous stem cell transplant (AuSCT) if categorized as favorable-risk or allogeneic stem cell transplant (ASCT) if adverse-risk. Intermediate-risk pts were to receive AuSCT or ASCT based on the post-consolidation levels of Measurable Residual Disease as measured by flow-cytometry.

Risk categorization was originally conducted according to NCCN2009 recommendations.

Among 500 patients, 445 (89%) were re-classified according to the ELN2017 criteria: ELN2017-FR (186/455; 41.8%), ELN2017-IR (179/445 40.2%) and ELN2017-AR (80/455; 18%); in 55 patients (11%) ELN2017 was not applicable (ELN2017-NC). Two-year overall survival (OS) was 68.8%, 51.3%, 45.8% and 42.8% for ELN2017-FR, ELN2017-IR, ELN2017-NC, and ELN2017-AR group, respectively (p<0.001). When comparing the two different transplant strategies in each ELN2017 risk category, a significant benefit of AuSCT over ASCT was observed among ELN2017-FR patients (2-years OS of 83.3% vs. 66.7%; p=0.0421). The two transplant procedures performed almost equally in the ELN2017-IR group (2-years OS of 73.9% vs. 70.8%; p=0.5552).

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.
INTRODUCTION

In 2010 and afterward in 2017, a panel of experts, convened on the behalf of the European LeukemiaNet (ELN), released and updated recommendations for harmonizing diagnosis, response and prognostic assessment of acute myeloid leukemia (AML).\(^1,2\) The latest version was substantially influenced by the evidence of the prognostic significance of the fms-related tyrosine kinase 3 (\textit{FLT3}) and Nucleophosmin 1 (\textit{NPM1}) gene mutation. Mutations of \textit{FLT3}, both at the tyrosine kinase domain (TKD) and the juxta-membrane level (ITD), are recognized in 20-25% of patients with a diagnosis of AML, thus representing the most common lesions found in this disease.\(^3,4\) In particular, FLT3-ITD mutation has historically been associated with poor prognosis due to an increased relapse rate and a dismal overall survival (OS). The ELN2017 recommendations have revisited and expanded this concept by capturing the prognostic meaning of FLT3-ITD allelic ratio\(^5,6\) and of concomitant mutations of FLT3-ITD and NPM1. Actually, mutations of \textit{NPM1} gene define a sub-group with a better outcome, with their presence being able to counterbalance the deleterious effect of the \textit{FLT3}-ITD mutations, when co-expressed.\(^7\) The interaction between NPM1 and \textit{FLT3-ITD} allelic ratio discriminates groups with different prognosis.\(^2\) Whereas \textit{NPM1} mutation (\textit{NPM1}\textsubscript{mut}) alone has retained its favorable prognostic role, the genotypes resulting from the combination of \textit{NPM1}\textsubscript{mut} and \textit{FLT3-ITD} with “low” (\textit{FLT3}\textsubscript{low}) or “high” (\textit{FLT3}\textsubscript{high}) allelic ratio are now distributed in all the 3 ELN2017 risk categories.\(^8,9\) Furthermore, the 2017 version of the ELN recommendations was further updated in that it included in the adverse group, several new gene mutations such as \textit{RUNX1}, \textit{ASXL1} and \textit{TP53}.\(^2\)
Recently, the Gruppo Italiano Malattie EMatologiche dell’Adulato (GIMEMA) Foundation accomplished a prospective, multicentric clinical trial (GIMEMA AML1310), strategy of which relied on the prognostic integration of pretreatment cytogenetics and genetics with post-consolidation MRD, as detected by multiparametric flow cytometry (MFC). Based on this strategy, patients were to receive post-consolidation autologous stem cell transplantation (AuSCT) or allogeneic stem cell transplantation (ASCT), respectively, depending on their risk profile.9–12 At the time when the study was designed, the National Comprehensive Cancer Network (NCCN) 2009 risk classification13 was the only available, therefore the patients were stratified according to this guideline. Accordingly, all patients carrying a FLT3 mutation were considered at high risk and allocated to ASCT, 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 been licensed yet, neither they were administered to patients enrolled in the trial.14

Five hundred patients were included in the final analysis of the AML1310 trial and the combination of baseline genetic with the assessment of the allelic ratio in FLT3-ITD positive (FLT3-ITDmut) cases allowed us to reclassify 445 of these 500 according to the new ELN2017 risk classification. At the moment of the present analysis, the mutational sequencing of RUNX1, ASXL1 and TP53 genes by NGS is not available yet.15

The aim of this 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 WHO diagnostic criteria\textsuperscript{16} were eligible for the GIMEMA AML1310 Study (\textit{EudraCT number 2010-023809-36; ClinicalTrials. Gov Identifier NCT01452646}). Main inclusion and exclusion criteria have been published elsewhere and are detailed in the Supplemental Material.\textsuperscript{9} 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 post remission therapy, intensity of which was risk-driven, prolonged the 2-years OS as compared to the historical data from the previous LAM99P GIMEMA trial.\textsuperscript{17} Upfront evaluation included bone marrow (BM) sampling for morphology, cytogenetics, molecular genetics and MFC analysis. The baseline MFC assessment was a necessary step to identify leukemia associated immunophenotypes (LAIP). Identification of baseline LAIPs by a high-sensitivity 8–color MFC assay was the essential requirement for monitoring measurable residual disease (MRD) after therapy, at the established post-consolidation time-point. Based on several retrospective validations in the context of former EORTC/GIMEMA protocols,\textsuperscript{12} the threshold for discriminating MRD negative from MRD positive cases was set at 3.5x10\textsuperscript{-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 \textit{NPM1}, \textit{FLT3-ITD}, \textit{FLT3-TKD}, \textit{c-KIT} and of rearrangements \textit{RUNX1-RUNX1T1} or \textit{CBFβ/MYH11}, defining core binding factor (\textit{CBF})-positive
AML. Using polymerase-chain reaction (PCR), the threshold discriminating low from high FLT3-ITD allelic ratio was set at 0.5, as proposed in the ELN2017 recommendations. 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. All molecular and MFC analyses were centralized at Laboratorio di Diagnostica Integrata Oncoematologica (Tor Vergata University Hospital, Rome, Italy), whereas baseline conventional karyotype was carried out at local institutions. BM and peripheral blood (PB) were used as sources to assess response to treatment, according to the recommendations of an international working group. 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 v.1 classification. Induction and consolidation regimens have been reported elsewhere and in the Supplemental Materials. By integrating the NCCN2009 classification with the level of post-consolidation MRD, 4 categories of risk were considered: favorable- (NCCN-FR) or poor-risk (NCCN-PR) patients, who were submitted to AuSCT or ASCT respectively; intermediate-MRD negative (NCCN-IR-Neg) or positive (NCCN-IR-Pos) patients, who were to receive AuSCT or ASCT, respectively. ASCT and AuSCT were to be performed within three months of the end of the consolidation course.

Statistical analysis and sample size calculation

The primary objective was the percentage of OS at two years. An estimated number of 213 subjects was 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 two
years is 50% and the alternative hypothesis that OS is 60%, using a Single-Stage Phase II design with a 5% significance level (based on data of the historic control group GIMEMA LAM99P). Based on the historical control group, we also considered that approximately 70% of the observed patients would have been classified as IR, therefore allowing to reach the figure of 150 patients available for MRD driven treatment allocation. However, after 173 subjects were enrolled, only 56 belonged to the IR category (32% vs 70% expected). Therefore, to reach the target of 150 subjects belonging to the IR category, an amendment to the protocol was adopted in 2013 and the sample size was adjusted to 515 subjects to recruit. Patients’ and disease characteristics were summarized by means of cross-tabulations for categorical variables or by quintiles for continuous variables. In univariate analysis, non-parametric tests were performed for comparisons among ELN2017 and combined NPM1/FLT3-ITD groups (Chi-Squared and Fisher Exact test in case of categorical variables or response rate, Mann-Whitney and Kruskal-Wallis test 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 using the Kaplan-Meier product limit estimator. Differences in terms of OS and DFS were evaluated by means of Log-Rank test in univariate analysis, and by means of Cox regression model in multivariate analysis, after assessment of proportionality of hazards. All variables with a p-value less than 0.15 in univariate analysis were considered into 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 were calculated at 95% level and all tests were two-sided, accepting p ≤0.05 as indicating a statistically significant difference. All analyses were
performed using the SAS (version 9.4) and R (R Foundation for Statistical Computing, Vienna, Austria) system software. Study data were collected and managed using the REDCap electronic data capture tools hosted at 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, therefore the final analysis included 500 individuals. Median age was 49 (18-61) years, with a slight male predominance (52%). Characteristics of the patients are summarized in Table 1.

Among 500 cases, FLT3-ITD and NPM1 mutations were detected in 123 (25.8%) and in 182 (38%) cases, respectively. Of 123 FLT3-ITD\textsuperscript{mut} and 182 NPM1\textsuperscript{mut}, co-expression was observed in 80 patients. FLT3-ITD allelic ratio, available in 109/123 (88.6%) patients, resulted to be low and high in 40 (32.5%) and 69 (56.1%) patients, respectively. The original, per-protocol risk stratification arranged patients as follows: 138 (27.6%) were NCCN-FR, 174 (34.8%) were NCCN-IR, and 188 (37.6%) were NCCN-PR. The 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 non-classifiable category named ELN2017-NC. Therefore, 445/500 (89%) patients were ELN2017 re-classified as follows: 186 (41.8%) were favorable risk (ELN2017-FR), 179 (40.2%) intermediate risk (ELN2017-IR), 80 (18%) adverse risk (ELN2017-AR). By adding up estimation of the allelic ratio, FLT3-ITD positive patients spread across all the 3 ELN2017 categories: 38 (31%) in the ELN2017-FR group, 51 (41%) in the ELN2017-IR group and 20 (16%) in the ELN2017-AR one. Four additional patients were moved from NCCN2009-PR 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 38/188 (20.2%), 55/188 (29.2%), and 20/188 (10.6%) NCCN2009-PR patients now being reclassified as ELN2017-FR, ELN2017-IR, and ELN2017-NC, respectively. Based on this, the analysis of the differential distribution of patients resulted in a 65.4% concordance (327/500) between the two risk stratification models. Indeed, only 75/188 (39%) 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 132/138 (95.6%) NCCN-FR and 120/174 (68.9%) NCCN-IR cases still being classified as ELN2017-FR and ELN2017-IR, respectively [Figure 1].

Overall, after a maximum of 2 cycles of induction, 361 (72%) patients obtained a CR/CRi: 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<0.001).

The majority of these patients (342/361, 95%) successfully started the consolidation phase and were subsequently allocated to the corresponding transplant procedure: 177 (52%) to AuSCT and 165 (48%) to ASCT. Of the 177 AuSCT candidates, 111 (62.7%) were transplanted (82 [73.9%] ELN2017-FR, 19 [17.1%] ELN2017-IR, 2 [1.8%] ELN2017-AR, 8 [7.2%] ELN2017-NC). The number of ASCT candidates, with the addition of 23 patients who achieved a CR after salvage therapy, was further enlarged to 188. Overall, out of 188 ASCT candidates, 132 (70.2%) were transplanted (25 [18.9%] ELN2017-FR, 61 [46.2%] ELN2017-IR, 25 [18.9%] ELN2017-AR, 21 [15.9%] ELN2017-NC) (p<0.001). Finally, 19 patients who were not submitted to AuSCT, neither ASCT, received additional HDARAC chemotherapy: 18 from the ELN2017-FR group, 1 from the ELN2017-IR one.

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, whereas CIR was 33% (95% CI 28-38). [Supplemental Figure 1S]

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 ones (2-year OS 42.8%). ELN2017-NC patients had an intermediate outcome, being their 2-year OS 45.8%. (p<0.001) [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=0.0297). [Figure 2B] Two-year CIR was 31.3%, 29.4%, 42.8%, and 39.2% for the ELN2017-FR, ELN2017-IR, ELN2017-AR and ELN2017-NC patients, respectively (p=0.2343). [Supplemental Figure 2S]

We next investigated the impact of the selected post-consolidation strategy (ASCT vs. AuSCT) on each ELN2017 category. Due to the AML1310 protocol design and the consequent negligible number of patients receiving AuSCT (only 2) in the ELN2017-AR group, comparison between the two 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 ASCT and those who received AuSCT (2-years OS of 70.8% and 73.9%, for ASCT and AuSCT, respectively; p=0.5552). On the other hand, belonging to the ELN2017-FR category and being submitted to AuSCT rather than to ASCT, was associated with a survival advantage (2-years OS of 83.3% vs. 66.7% for AuSCT and ASCT, respectively; p=0.0421). [Figure 3A-B]. Among ELN2017-IR category, non-relapse mortality (NRM) for patients submitted to AuSCT or ASCT was 5.3% and 13.7% (p=0.3246), respectively. In the
ELN2017-FR category, NRM was 0% and 16.7% (p=0.0002), for those receiving AuSCT or ASCT, respectively.

Considering that in the original AML1310 protocol the post-consolidation treatment of the NCCN2009-IR category was decided according to the MRD status, we analyzed if MRD maintains its role also in the new designed ELN2017-IR category. Post consolidation MRD was available in 85 patients, 55 negative (64.7%) and 30 positive (36.3%). In the overall population MRD negative patients had a better 2-years OS (76.5 vs 58.8) although not significant (p=0.247). Among MRD negative patients, 26 underwent ASCT and 10 AuSCT, respectively, whereas among MRD positive patients 17 underwent ASCT and 3 AuSCT, respectively. The 2-years OS, when landmarked “from” transplant, was not different between AuSCT and ASCT in MRD negative patients (85.7 vs. 77.8, p=0.234) whereas, among MRD positive ones, was significantly longer for those receiving ASCT (75% vs 0, p=0.0231) (Supplemental Figures 3S).

Univariate analyses evidenced the independent role of each considered covariate in influencing duration of OS. Indeed, as compared to the ELN2017-FR group, belonging to the ELN2017-AR (HR=2.203, CI 1.496-3.246; p<0.0001), ELN2017-IR (HR=1.796, CI 1.293-2.494; p=0.0005), and ELN2017-NC (HR=2.267, CI 1.488-3.455; p=0.0001) category was associated with a shorter duration of OS.

In the multivariate model for OS prediction, it was confirmed the significant prognostic role of age (see also Supplemental Figure 4S), ELN2017 risk stratification and transplantation as a time-dependent parameter [Table 2].
DISCUSSION

In our analysis we demonstrated that, when implemented in the prospective, risk-adapted, MRD-driven, AML1310 GIMEMA protocol, ELN2017 risk classification maintained its prognostic significance, therefore impacting on OS and DFS.

Prognostic stratification of AML has been substantially improved by the introduction of ELN2017 recommendations. Indeed, through the combination of a broad range of genetic and cytogenetic abnormalities, the ELN2017 recommendations identify three prognostic classes (favorable, intermediate and adverse). The ELN2017 version, as compared to the precedent ELN2010 edition, is influenced by the evidence that the mutual interactions between NPM1 and FLT3, and the FLT3-ITD allelic ratio play a crucial role in discriminating patients with different outcome. Based on this, the 2017 revisited edition has resulted in a redistribution of FLT3-ITD mutations across all the three risk categories: cases $NPM1^{\text{mut}}/FLT3^{\text{ITD}^{\text{low}}}$ have been allocated in the favorable group, those $NPM1^{\text{mut}}/FLT3^{\text{ITD}^{\text{high}}}$ or $NPM1^{\text{wt}}/FLT3^{\text{ITD}^{\text{low}}}$ in the intermediate group, and finally cases $NPM1^{\text{wt}}/FLT3^{\text{ITD}^{\text{high}}}$ in the adverse one. The reliability of this risk stratification has been widely recognized, leading to its incorporation into the NCCN clinical practice guidelines for AML.22

Patients recruited to the GIMEMA AML1310 trial were categorized based on the NCCN2009 criteria.13 Therefore, the purpose of this post-hoc analysis aimed at verifying if the prospective, risk-adapted strategy of GIMEMA AML1310 trial still held true when the same population was retrospectively re-arranged according to the ELN2017 classification.

The exercise of a retrospective application of ELN2017 criteria to the AML1310 cohort was feasible in approximately 90% of the patients. As consequence of the new role assigned to
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 excluded the small group of 55 ELN2017-NC patients (11% of the whole cohort), 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 three 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%, as compared to 69.9% of NCCN-PR. The CR rate achieved by the category of ELN2017-AR patients appears more logically connected to the risk profile of this group than the CR frequency demonstrated for the NCCN2009-PR group. This observation seems even more robust if one considers the prospective nature of 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 ASCT or AuSCT. In the same line, also the ELN2017-IR category included a composite population of patients since, 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, AuSCT was associated with a better
2-years OS as compared to ASCT. We assume that the favorable outcome of patients within the ELN2017-FR category was jeopardized by the NRM of ASCT. On the other hand, in the ELN2017-IR category, AuSCT and ASCT 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 ASCT counterbalances the poor prognosis of a MRD positive status and prolongs OS of MRD positive patients to equalize the one of MRD negative ones. This confirms that post-consolidation MRD assessment, whatever 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 approximately 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. Since almost half of these patients (26/55, 47%) were not submitted to either AuSCT or ASCT, we speculate that their dismal outcome could be attributed to the delivery of a sub-optimal post-consolidation therapy.

We are aware of the limitations of our study, mainly attributable to the post-hoc nature of this analysis. The GIMEMA AML1310 trial was not designed for the purpose of the current analysis, and at the moment of study design not even the 2010 version of the ELN recommendations was available. Further, 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 NGS analysis of the present series is currently ongoing. Unfortunately, at the moment of the present analysis, results are not available yet.

In conclusion, our attempt to re-classify, 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 the outcome. In the context of the GIMEMA AML1310 trial, these patients received homogeneous induction and consolidation courses, with the post-consolidation 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 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 present 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 enhance the outcome prediction of ELN2017 classification.
Data Sharing Statement

For original AML1310 trial collection data methods please contact adriano.venditti@uniroma2.it and refer to EudraCT #2010-023809-36; www.clinicaltrials.gov Identifier #NCT01452646.

Author Contributions and Disclosures

F.B., R.P., A.V., A.P., and V.A., created the concept and design; A.C., L.Me, V. C., R.C., P. D.F., G.S., P.S., F.L., G.M., M.L., S.C., L.Ma., 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; F.B., L.M., 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.

F.B. and R.P. equally contributed to the manuscript

Conflict of Interest Statement

None of the authors has a relevant conflict of interest.
REFERENCES

1. Döhner H, Estey EHE, Amadori S, et al. 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):1079–89.

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. Differential impact of allelic ratio and insertion site in FLT3-ITD-positive AML with respect to allogeneic transplantation. *Blood*. 2014;124(23):3441–9.

6. Small D. FLT3 mutations: biology and treatment. *Hematology Am. Soc. Hematol. Educ. Program*. 2006;178–184.

7. Ivey A, Hills RK, Simpson MA, et al. 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. 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. *Survival (Lond)*. 2010;116(13):2295–2303.

11. Buccisano F, Maurillo L, Del Principe ML, 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. Acute Myeloid Leukemia, Version 1.2009, NCCN Clinical Practice Guidelines in Oncology. 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. 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, Steudel 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. 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. *J. Clin. Oncol.* 2003;21(24):4642–4649.

22. Tallman MS, Wang ES, Altman JK, et al. Acute Myeloid Leukemia, Version 3.2019, NCCN Clinical Practice Guidelines in Oncology. *J. Natl. Compr. Canc. Netw.* 2019;17(6):721–749.
Table 1. General characteristic of the study population

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

Abbreviations: FR, favorable risk; IR, intermediate risk; AR, adverse risk; FLT-ITD<sup>mut</sup>, FLT3-ITD mutated; FLT3<sup>low</sup>, FLT3-ITD mutated with low allelic ratio; FLT3<sup>high</sup>, FLT3-ITD mutated with high allelic ratio; NPM1<sup>mut</sup>, NPM1 mutated; NCCN, National Comprehensive Cancer Network risk stratification; PR, poor risk; ELN2017, European Leukemia Net risk stratification version 2017; NC, not classifiable.
Table 2. Multivariate model for OS prediction

| Parameter               | Detail                                      | Probability (ChiSquare) | Hazard Ratio | 95% Lower Confidence interval | 95% Upper Confidence interval |
|-------------------------|---------------------------------------------|-------------------------|--------------|------------------------------|------------------------------|
| Age                     |                                             | <.0001                  | 1.033        | 1.019                        | 1.048                        |
| ELN2017 risk group      | ELN2017-NC vs ELN2017-FR                    | 0.0003                  | 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                    | 0.0003                  | 1.838        | 1.321                        | 2.557                        |
| Transplant, covariate   | transplant vs. no transplant                | 0.0185                  | 0.674        | 0.485                        | 0.936                        |
| time-dependent          |                                             |                         |              |                              |                              |

Abbreviations: ELN2017, European Leukemia Net risk stratification version 2017; NC, not classifiable; FR, favorable risk; IR, intermediate risk; AR, adverse risk.
Figure 1. Spine plot for each NCCN showing the proportion of Adverse (dark gray), Intermediate (medium gray), and Favorable (light gray). A small proportion of NCCN2009 patients (75/188 [39%]) were still classified as high risk also according to ELN2017. A high proportion of NCCN2009-PR cases were redistributed across all ELN2017 risk groups with 38/188 (20.2%), 55/188 (29.2%), and 20/188 (10.6%) NCCN2009-PR patients now being reclassified as ELN2017-FR, ELN2017-IR, and ELN2017-NC, respectively. At variance, 132/138 (95.6%) NCCN-FR and 120/174 (68.9%) NCCN-IR cases were still classified as ELN2017-FR and ELN2017-IR, respectively.

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

Figure 3. Correlation between post-consolidation strategy and outcome for each ELN2017 risk category. (A) Benefit of AuSCT in the ELN2017-FR category (2-years OS of 83.3% vs. 66.7% for AuSCT vs. AlloSCT). (B) Almost equal performance of AlloSCT and AuSCT in the ELN2017-IR category (2-years OS of 70.8%, 73.9%, for AlloSCT and respectively).
Figure 1. Spine plot for each NCCN showing the proportion of Adverse (dark gray), Intermediate (medium gray), and Favorable (light gray). A small proportion of NCCN2009 patients (75/188 [39%]) were still classified as high risk also according to ELN2017. A high proportion of NCCN2009-PR cases were redistributed across all ELN2017 risk groups with 38/188 (20.2%), 55/188 (29.2%), and 20/188 (10.6%) NCCN2009-PR patients now being reclassified as ELN2017-FR, ELN2017-IR, and ELN2017-NC, respectively. At variance, 132/138 (95.6%) NCCN-IR and 120/174 (68.9%) NCCN-IR cases were still classified as ELN2017-FR and ELN2017-IR, respectively.

Abbreviations: ELN2017, European Leukemia Net risk stratification version 2017; NCCN 2009, National Comprehensive Cancer Network risk stratification version 2009; PR, poor risk; IR, intermediate risk; FR, favorable risk.

Figure Legend: Dark grey, ELN2017 adverse risk; Medium grey, ELN2017 intermediate risk; Light grey, ELN2017 favorable risk.
Figure 2. Patients’ outcome according to ELN2017 risk stratification and FLT3/NPM1 genes interactions. (A) Two-years OS was 68.8%, 51.3%, 42.8% and 45.8% for patients belonging to ELN2017-FR, ELN2017-IR, ELN2017-AR, and ELN2017-NC categories, respectively. (B) Two-years DFS was 59.9%, 54.2%, 45.5%, and 40.3% for the ELN2017-FR, ELN2017-IR, ELN2017-AR and ELN2017-NC patients, respectively.
Figure 3. Correlation between post-consolidation strategy and outcome for each ELN2017 risk category. (A) Benefit of AuSCT in the ELN2017-FR category (2-years OS of 83.3% vs. 66.7% for AuSCT vs. AlloSCT). (B) Almost equal performance of AlloSCT and AuSCT in the ELN2017-IR category (2-years OS of 70.8%, 73.9%, for AlloSCT and respectively).

A. ELN2017-FR OS

B. ELN2017-IR OS