haematologica
Journal of The Ferrata Storti Foundation

Clinical relevance of an objective - limit of detection - limit of quantification - based flow cytometry approach for measurable residual disease assessment in acute myeloid leukemia. A post-hoc analysis of the GIMEMA AML1310 trial

by Francesco Buccisano, Raffaele Palmieri, Alfonso Piciocchi, Valentina Arena, Luca Maurillo, Maria-Ilaria Del Principe, Giovangiacinto Paterno, Maria-Antonietta Irno Consalvo, Tiziana Ottone, Mariadomenica Divona, Consuelo Conti, Daniela Fraboni, Serena Lavorgna, William Arcese, Maria Teresa Voso, and Adriano Venditti

Received: August 1, 2021.
Accepted: February 9, 2022.

Citation: Francesco Buccisano, Raffaele Palmieri, Alfonso Piciocchi, Valentina Arena, Luca Maurillo, Maria-Ilaria Del Principe, Giovangiacinto Paterno, Maria-Antonietta Irno Consalvo, Tiziana Ottone, Mariadomenica Divona, Consuelo Conti, Daniela Fraboni, Serena Lavorgna, William Arcese, Maria Teresa Voso, and Adriano Venditti. Clinical relevance of an objective - limit of detection - limit of quantification - based flow cytometry approach for measurable residual disease assessment in acute myeloid leukemia. A post-hoc analysis of the GIMEMA AML1310 trial. Haematologica. 2022 Mar 17. doi: 10.3324/haematol.2021.279777. [Epub ahead of print]

Clinical relevance of an objective – limit of detection – limit of quantification - based flow cytometry approach for measurable residual disease assessment in acute myeloid leukemia. A post-hoc analysis of the GIMEMA AML1310 trial.

Francesco Buccisano[1†], Raffaele Palmieri[1†], Alfonso Piciocchi[2], Valentina Arena,[2] Luca Maurillo[1], Maria-Ilaria Del Principe[1], Giovangiacinto Paterno[1], Maria-Antonietta Irno-Consalvo[1], Tiziana Ottone[1], Mariadomenica Divona[1], Consuelo Conti[1], Daniela Fraboni[1], Serena Lavorgna[1], William Arcese[1,3], Maria Teresa Voso[1], Adriano Venditti[1]

1 Ematologia, Dipartimento di Biomedicina e Prevenzione, Università di Roma "Tor Vergata", Roma,
2 Centro Dati Fondazione GIMEMA Onlus, Roma, Italia;
3 Rome Transplant Network, Roma, Italy

[†] The authors contributed equally to the manuscript

Keyword: Acute Myeloid Leukemia; Minimal Residual Disease; Immunophenotyping; Measurable Residual
Disease

WORDS: 3293
FIGURES: 3
TABLES: 3
REFERENCES: 36

CONFLICT OF INTEREST
The authors declare no conflict of interest.

Author contributions FB, RP, AV designed the study, FB, AP and VA collected and analyzed the data, FB, RP, AP, VA, LM, MIDP, GP, MAIC, TO, MD, CC, DF, SL, WA, MTV and AV wrote the paper and approved its final version.

ACKNOWLEDGEMENTS
The authors are indebted to Prof. Bruno Brando and Dr. Arianna Gatti for the critical review of the manuscript and the helpful suggestions.

Data sharing statement: For original, anonymized data, please contact the corresponding author (francesco.buccisano@uniroma2.it).

Correspondence:
Francesco Buccisano, MD
Department of Biomedicine and Prevention
Tor Vergata University of Rome
Mail: francesco.buccisano@uniroma2.it; Tel.: +39-06-20903228

Abstract

Using a multiparametric flow cytometry (MFC) assay, we assessed the predictive power of a threshold calculated applying the criteria of limit of detection (LOD) and limit of quantitation (LOQ) in adult patients affected with Acute Myeloid Leukemia (AML). This was a post-hoc analysis of 261 patients enrolled in the GIMEMA AML1310 prospective trial. According to the protocol design, using the predefined MRD threshold of 0.035% bone marrow residual leukemic cell (RLC) calculated on mononuclear cells, 154 (59%) were negative (MRD<0.035%) and 107 (41%) were positive (MRD≥0.035%). Using LOD and LOQ, we selected the following categories of patients: 1) LOD^{neg} if RLC were below LOD (74; 28.4%); 2) LOD^{pos}-LOQ^{neg} if RLC were between LOD and LOQ (43; 16.5%); and 3) LOQ^{pos} if RLC were above LOQ (144; 54.4%). Two-year overall survival (OS) of these 3 categories was 75.4% vs. 79.8% vs. 66.4%, respectively (p=0.1197). Due to superimposable outcome, LOD^{neg} and LOD^{pos}-LOQ^{neg} categories were combined. Two-year OS of LOD^{neg}/LOD^{pos}-LOQ^{neg} patients was 77.0% versus 66.4% of LOQ^{pos} individuals (P=0.043). Such a figure was challenged in multivariate analysis (p=0.048, HR 0.628, 95% CI 0.396-0.997) that confirmed the independent role of LOD-LOQ approach in influencing OS. In the AML1310 protocol, using the threshold of 0.035%, 2-year OS of MRD<0.035% and MRD≥0.035% patients was 74.5% vs. 66.4%, respectively (p=0.3521). In conclusion, the use of LOD-LOQ method results in a more sensitive detection of MRD that, in turn, translates in a more accurate recognition of patients with different outcome.

Introduction

Measurable residual disease (MRD) is increasingly employed as a biomarker of quality of complete remission (CR) in AML patients treated by intensive chemotherapy.[1] Multiparametric Flow Cytometry (MFC), together with Reverse Transcriptase quantitative Polymerase Chain Reaction (RT-qPCR) are the two leading techniques for MRD quantification. Recent studies indicate that, due to technical improvements and the availability of up to 8-10 color immunostainings, MFC specificity and sensitivity may be reliably increased, provided that a sufficient number of relevant events is acquired.[2,3] In B-cell precursor acute lymphoid leukemia (BCP-ALL) and multiple myeloma (MM), the use of standardized panels and the acquisition of large numbers of events (>4 millions) led to MRD assessment by MFC becoming at least as sensitive as PCR-based methods.[4,5] Likewise, sensitivity of MRD determination in MM and chronic lymphocytic leukemia (CLL) improved dramatically up to 10^{-5}-10^{-6} as soon as larger numbers of events (3-5 millions) were acquired, in the context of so-called next-generation flow (NGF).[6–8]

In AML, the number of clustered events and the denominator of acquired events necessary for a reliable MRD recognition are poorly standardized and may be affected by several technical and clinical variables. In myeloid bone marrow (BM), particularly in regenerating phases after chemotherapy, the normal maturational patterns may interfere with the detection of leukemia-associated immunophenotypes (LAIPs) generating a relevant background noise. Likewise, although at a lesser extent, this background noise may affect the identification of the putative "empty spaces" when MRD is detected by a "different-from-normal" approach.[9]

The consensus of the European LeukemiaNet (ELN) MRD working party suggests that a MRD threshold of 0.1% is informative for clinical decisions once that 500'000-1'000'000 events are acquired.[10] Such a target of acquired events guarantees that the threshold of 0.1% has a reliable sensitivity and a sufficient specificity, because no LAIPs have been detected above this threshold even in regenerating BM.[11] Nonetheless, the same guidelines suggest that MRD tests with MRD quantified below 0.1% may still be consistent with residual leukemia, indeed several studies have shown prognostic significance of MRD levels below 0.1%.[12–16]

In the GIMEMA AML1310 protocol, post-remission therapy of young patients with AML was decided combining cytogenetics/genetics and post-consolidation levels of MRD after consolidation cycle as measured by MFC.[17] Intermediate-risk patients were to receive autologous stem cell transplant (AuSCT) or allogeneic stem cell transplant (ASCT) depending on the post-consolidation levels of MRD. The threshold of negativity was set below 0.035% residual leukemic cells (RLC) as measured on mononuclear cells (MNC). This threshold was selected after repeatedly validating it in retrospective, sequential cohorts of patients enrolled in former EORTC/GIMEMA protocols AML10, AML12, AML13, AML15 and AML17.[18–21] In the AML1310 protocol we have confirmed, prospectively, that the threshold of 0.035% retains as the same predictive value as in the retrospective analyses.[16]

However, since the previous EORTC/GIMEMA and AML1310 protocols have in common the same therapeutic schedule either in induction or in consolidation, one can argue that the threshold of 0.035% may be protocol specific so that it cannot be applied universally. In fact, thresholds in AML are often selected retrospectively based on their association with outcomes. Accordingly, confirmatory, prospective validations are required.[22,23] In the attempt to overcome such a "protocol-effect" and to reliably improve the statistical accuracy of MRD assessment, we revised the post-consolidation MRD determinations of the GIMEMA AML1310 protocol by calculating, for each case, limit of detection (LOD) and limit of quantification (LOQ). As in MM and CLL, the target of 20 and 50 relevant events in the final gate, respectively, were adopted. According to the ELN guidelines, the analysis was conducted on CD45 expressing elements.[3,24] Patients were classified as negative (LOD^{neg}), positive not quantifiable (LOD^{pos}-LOQ^{neg}) and positive

quantifiable (LOQ^{pos}). Due to the retrospective nature of the analysis, we were not able to establish a limit of blank (LOB) to properly exclude the background noise of each aberrant phenotype selected for MRD assessment.

In our exploratory analysis, the new MRD categories were compared to the protocol reference threshold of 0.035%, the genetic/cytogenetic subgroups and the post-remission treatments. To the best of our knowledge, this is the first time that an absolute threshold based on LOD and LOQ is applied to assess MRD in AML by MFC. We believe that the analysis of a prospective series of homogeneously treated patients, represents a unique chance to corroborate the robustness of LOD and LOQ approach in MRD determination in AML.

METHODS

Patients

Previously untreated patients with a diagnosis of *de novo* AML according to the WHO diagnostic criteria[25] were eligible for the GIMEMA AML1310 Study (*EudraCT number 2010-023809-36; ClinicalTrials. Gov Identifier NCT01452646*) (Supplemental methods). [16,26] The present analysis has been performed with different purposes on a subgroup of 263 patients whose MRD was determined after consolidation cycle. The study was approved by the ethical committees of the participating Hospitals/Academic Institutions and was conducted in accordance with the Declaration of Helsinki. All participants gave their informed consent.

LOD and LOQ calculation

There are numerous studies demonstrating that 20 events are a conservative value for the smallest (homogeneous) population that can be detected in a given flow cytometric list mode data file by experienced operators. This implies that the LOD can be estimated as (20/total number of cells analyzed) × 100%.[26] Similarly, it is also widely accepted that more than 50 events is a reasonable threshold for reproducible enumeration of a cell population by experienced operators; consequently, the LOQ can be estimated as (50/total number of cells analyzed) × 100%.[27] Thus, the LOD and the LOQ will be both typically dependent on the total number of cells analyzed. LOD and LOQ were established at 20 and 50 clustering events expressing a LAIP, respectively, and counted on CD45 expressing events according to the ELN recommendations.[10] Based on such an approach, patients were classified as MRD negative if RLC were below LOD (LOD^{neg}), MRD positive non quantifiable if RLC were between LOD and LOQ (LOD^{pos}-LOQ^{neg}) and MRD positive quantifiable if RLC were above LOQ (LOQ^{pos}). Samples were acquired by a FacSCanto II (Becton Dickinson, Mountain View, CA, USA). Data analysis was performed using Infinicyt-software version 1.7 (Cytognos SL, Salamanca, Spain).

Statistical analysis

OS (time elapsed from treatment start to death) and DFS (time from 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. Cumulative incidence of relapse (CIR) was estimated by cumulative incidence curves using the proper non-parametric method. Patients' and disease characteristics were summarized by means of cross-tabulations for categorical variables or by quintiles for continuous variables. Differences between categorical variables or response rates in subgroups were tested by the chi-squared or Fisher exact tests, as appropriate. Confidence intervals were calculated at 95% level and all tests were two-sided, accepting $p \leq 0{\cdot}05$ as indicating a statistically significant difference. All covariates were evaluated in univariate models and all factors with univariate association within p-value <0.1 were considered in the multivariate models as potential parameters. Backward and stepwise methods were applied to identify the multivariate models with a step-by-step iterative construction that involves the selection of independent variables to be considered in the final model. 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 REDCap20 electronic data capture tools hosted at GIMEMA Foundation.

RESULTS

The present analysis includes 261 patients in whom a post-consolidation BM sample was collected and sent to the central laboratory for MRD determination. Clinical characteristics of the patients are summarized in **Table 1**. Subjects with a percentage of RLC equal or above 0.035% of the total number of acquired MNC qualified as $MRD^{\geq 0.035\%}$. In the same 261 patients, LOD and LOQ were calculated on CD45 expressing elements. Median number of MNC acquired was 559'197 (range 100'450-1'561'221) and the median number of CD45 expressing cells was 538'527 (range 88'040-1'548'172). Overall, 74/261 cases (28.4%) were classified as LOD^{neg}, whereas 144/261 (55.2%) and 43/261 (16.5%) were classified as LOD^{pos}-LOQ^{neg} and LOQ^{pos}, respectively (**Table 1S**). The target of 500'000 processed CD45+ events was reached in 158/261 (60.5%). The calculated median LOD and LOQ values were 0.0037 (0.0013-0.0227) and 0.0093 (0.0032-0.0568) (**Table 1S**).

According to the protocol MRD threshold of 0.035%, 107/261 (41.0%) were $MRD^{\geq 0.035\%}$ and 154/261 (59.0%) $MRD^{<0.035\%}$, respectively. The interactions between the different MRD estimates are summarized in **Table 2S**. Overall, 105/107 (98.1%) $MRD^{\geq 0.035\%}$ patients were LOQ^{pos} whereas only 74/154 (48.1%) $MRD^{<0.035\%}$ ones were LOD^{neg}-LOQ^{neg} ($p<0.001$). In fact, 41 (26.6%) and 39 (25.3%) of 154 $MRD^{<0.035\%}$ patients were reclassified as LOD^{pos}-LOQ^{neg} and LOQ^{pos}, respectively.

On the whole population, 2-year OS and DFS were 71.2% and 57.5%, respectively. No difference was observed in duration of OS between $MRD^{<0.035\%}$ and $MRD^{\geq 0.035\%}$ patients (74.5% vs. 66.4%, p=0.3521, **Figure 1A**). When the survival analysis was conducted according to the new categories that we identified, patients who were LOD^{neg} or LOD^{pos}-LOQ^{neg} had a superior OS as compared to LOQ^{pos} (75.4% and 79.8% vs. 66.4%), although the difference was not statistically significant (p=0.119). The equivalent outcome of LOD^{neg} and LOD^{pos}-LOQ^{neg} patients (**Figure1B**) persuaded us to aggregate these subgroups. Accordingly, we sorted 2 categories of patients, (LOD^{neg}/LOD^{pos}-LOQ^{neg}) and LOQ^{pos} whose duration of OS was statistically different (77.0% vs. 66.4%, p=0.0437), as depicted in **Figure 1C**.

As a further step of investigation, we repeated our analysis on 158/261 (60.5%) patients in whom ≥500'000 CD45+ events were acquired. This was to test whether a more numerically robust denominator enhanced specificity and then prognostic power of LOD-LOQ estimate. The threshold-based MRD allocation ($MRD^{<0.035\%}$ 82.4% vs. 67.2% of $MRD^{\geq 0.035\%}$, p=0.064, **Figure 2A**) was less effective in discriminating 2-year OS, whereas 2-year OS of LOD^{neg} and LOD^{pos}-LOQ^{neg} was superior than LOQ^{pos} (82.1% and 95.7% vs. 69.0%, p=0.014) with a significant difference between LOQ^{pos} and both LOD^{neg} and LOD^{pos}-LOQ^{neg} (p=0.038 and p=0.024, respectively, **Figure 2B**). LOD^{neg}/LOD^{pos}-LOQ^{neg} category resulted in a subset of patients with a strongly favorable outcome as compared to LOQ^{pos} subgroup (2-year OS of 86.7% vs. 69.0%, p=0.004, **Figure 2C**).

Then, we tried to integrate the MRD and LOD-LOQ model. By doing so, we generated 3 categories of patients ($MRD^{<0.035\%}LOD^{neg}$/LOD^{pos}-LOQ^{neg}, $MRD^{<0.035\%}LOQ^{pos}$, $MRD^{\geq 0.035\%}LOQ^{pos}$), whose features are shown in **Table 2**. A fourth category ($MRD^{\geq 0.035\%}LOD^{neg}$/$LOD^{pos}$-$LOQ^{neg}$) was dropped from the analysis because represented by only 2 patients.

Notably, $MRD^{<0.035\%}LOD^{neg}$/LOD^{pos}-LOQ^{neg} had a longer 2-year OS not only when compared to $MRD^{\geq 0.035\%}LOQ^{pos}$ but also when compared to $MRD^{<0.035\%}LOQ^{pos}$ patients, whose median MRD percentage was 0.016% (range 0.006-0.032). This comparison did not reach a statistical significance when the overall series was analyzed (76.7% vs. 67.5% and 65.9%, p=0.116) but was clearly significant when patients with at least 500'000 events were taken into account. More in detail, for patients in whom ≥500'000 CD45 expressing events were acquired, those $MRD^{<0.035\%}LOD^{neg}$/LOD^{pos}-LOQ^{neg} had a longer duration of OS as compared to $MRD^{<0.035\%}LOQ^{pos}$ and $MRD^{\geq 0.035\%}LOQ^{pos}$ (86.7%, 72.5% and 67.0%, respectively, p=0.018). Even more, $MRD^{<0.035\%}$ patients had a statistically different OS if they tested LOD^{neg}/LOD^{pos}-LOQ^{neg} or LOQ^{pos} (86.7% vs 72.5%, p=0.007) (**Figure 3**). To avoid a possible bias deriving from the original design of the

protocol, where MRD was used to address treatment only in the intermediate-risk category, we conducted the same analysis in the 77 patients belonging to this category. The results (**Figure 1S**) were completely superimposable (p=0.0286).

Finally, we explored the interaction of LOD^{neg}, LOD^{pos}-LOQ^{neg} and LOQ^{pos} categories with the post-remission treatment received (AuSCT, ASCT and no graft). As shown in **Figure 2S**, LOD^{neg}/LOD^{pos}-LOQ^{neg} patients submitted to AuSCT had the best 2-year OS (88.9%) as compared to the other categories (p=0.026). Notably, these patients, benefitted the most from AuSCT (88.9%) than from no-graft (55.9%, p=0.017) or ASCT (76.5%, p=0.089).

All clinical variables testing significant in univariate analysis were challenged in multivariate model (**Table 3**). The analysis confirms, in multivariate analysis, the independent impact on OS of poor-risk upfront classification (p=0.046, HR 0.625, 95% CI 0.394-0.991), ASCT (p=0.005, HR 0.47, 95% CI 0.28-0.80) and $MRD^{<0.035\%}LOQ^{pos}$ status (p=0.021, HR 2.19, 95% CI 1.13-4.27). Multivariate model including LOD-LOQ stratification and transplant as a time-dependent covariate resulted in achievement of significant p either in univariate (p<0.001, HR 5.02, 95% CI 2.31-10.9) or multivariable analysis (p=0.048, HR 0.628, 95% CI 0.396-0.997) for LOD-LOQ stratification whereas graft did not.

DISCUSSION

In this preliminary report, we demonstrated that an MRD estimate based on LOD and LOQ of CD45 expressing cells predicts survival of AML patients more accurately than the pre-established threshold of 0.035% RLC of MNC, which was used in the AML1310 protocol. Moreover, we observed that the predictive power of the LOD-LOQ approach increases proportionally with the number of events acquired (higher or lower than 500'000).

The search for the most informative value of MRD for clinical purpose remains a matter of debate in AML. The general experience indicates that many technical, biological and clinical confounding factors interfere with the identification of the "absolute threshold" below or above which the prognosis is more accurately predicted.[28] In fact, the background noise due to the normal maturational curves of BM precursors has forced researchers to define the MRD status as above or below a given level, which is able to anticipate a different clinical outcome, rather than as negative or positive.[23,29] Finally, the multifaceted interpretation of MRD is made even more complicated as a consequence of the therapy delivered. Different treatment schedules may have different thresholds of prognostic significance. These thresholds are currently selected by different approaches, in some cases applying empirical logarithmic scales or quartile segregation, in others applying specific statistical methods (e.g. ROC analysis or maximally-selected log-rank statistics).[23] A comprehensive review of the literature [30] prompted ELN panel to recommend a threshold of 0.1% not because it was the most predictive but because it was used and found relevant in the majority of the published studies.[10] Nevertheless, the panel of experts was well aware that levels of MRD below 0.1% are consistent with residual leukemia and that further efforts should be made to identify and validate lower thresholds. In theory, the validation of MRD as a clinical biomarker should rely on the well-designed analysis of retrospective case series, leading to the identification of informative thresholds. Afterwards, these thresholds are to be validated in prospective, MRD-oriented trials.[22] Despite this attempts, doubts will still persist because of the many different therapeutic contexts that can hamper the universal applicability of the selected thresholds. Indeed, the last FDA MRD guidance for the development of novel agents, raised concerns about the role of MRD as a surrogate end-point. Such concerns were due to the biological heterogeneity of AML and the lack of prospective studies having MRD negativity as a primary endpoint.[31,32] Furthermore, the putative threshold of sensitivity of the MRD assay should be at least 10-fold (1-log) below the clinical decision-making threshold.[31]

At variance, in other pathologies (e.g. ALL, MM and CLL), MRD assessment by MFC is highly standardized and reproducible in different treatment scenarios, so that it is proposed as a surrogate end-point in clinical trials.[31] In these diseases, an innovative approach called next-generation flow (NGF) has substantially improved the performance of standard MFC that now reaches levels of sensitivity comparable to those of RT-qPCR (10^{-4}–10^{-6}).[4–6,27] Such an approach requires the application of a minimum of 8-color panel and the acquisition of several millions of relevant events.[4,24] Using this approach in CLL, it was demonstrated that an MRD threshold of 0.01% (10^{-4}) was an independent predictor of progression free survival in patients treated with either chemo immunotherapy or novel agents.[33]

In the GIMEMA AML1310 trial, patients with intermediate risk defined according to the NCCN 2009[34] were addressed to ASCT or AuSCT if MRD positive or negative, respectively, after the consolidation cycle.[17] The threshold defining MRD negativity (0.035%) was validated in several retrospective analyses of previous EORTC/GIMEMA trials. In those analyses the threshold of 0.035% allowed to discriminate patients with a well distinct long-term prognosis across different genetic/cytogenetic subgroups.[14,18–20] This threshold was prospectively validated in the AML1310 trial, in which delivery of ASCT prolonged the OS of MRD positive intermediate-risk patients to equalize the one of MRD negative intermediate-risk patients, who received AuSCT.[16]

The working hypothesis leading to the current analysis was that an MRD estimate based on LOD-LOQ approach might further refine the outcome prediction of the 0.035% threshold. In AML1310 trial, we assumed that the MRD-oriented post-remission strategy (ASCT vs AuSCT) used for patients belonging to the intermediate-risk category, nullified the poor prognostic weight of MRD positivity. This resulted in an equivalent duration of OS and DFS of MRD negative and MRD positive patients, with MRD positivity losing its independent prognostic value in multivariate analysis, as compared to genetic-cytogenetic risk and post-remission treatment.[16] In contrast, the LOD-LOQ calculation of MRD discriminated two populations of patients (LOD^{neg}/LOD^{pos}-LOQ^{neg} and LOQ^{pos}) with a statistically significant different duration of OS. The multivariable analysis confirmed the independent prognostic role of the LOD-LOQ approach.

The power of the LOD-LOQ outcome prediction increased when the analysis included only samples in which the count of CD45 expressing events was at least 500'000. This observation confirms that, when dealing with rare events identification, the larger the denominator of relevant events the more accurate the target population estimation, provided that an adequate number of relevant events is collected (i.e. 20 for LOD and 50 for LOQ calculations). Furthermore, the availability of a marker allowing an easier extrapolation of the cell population in study (e.g. CD45) increases the accuracy of the measurement. This has also been proven true by others when MRD was determined only on the population defined by immature markers.[35]

Based on this, we assume that the LOD-LOQ MRD estimate is more accurate than the $MRD^{0.035\%}$ threshold for it allowed a superior discrimination within the $MRD^{<0.035\%}$ category. In fact, among $MRD^{<0.035\%}$ patients, the LOD^{neg} and LOD^{pos}-LOQ^{neg} status identifies "true negative" or "non quantifiable" cases with a better outcome. These patients might have been cured of their disease without ASCT, as demonstrated in a further subgroup analysis where patients submitted to AuSCT showed a very favorable outcome (**Figure 2S**). Interestingly, in our hands, LOD^{neg} and LOD^{pos}-LOQ^{neg} patients showed the same OS. We hypothesized at least 2 technical explanations. First, the median number of CD45+ event acquired may be not sufficient. In fact, the category of LOD^{pos}-LOQ^{neg} patients might be progressively narrowed if a very high number of relevant events is acquired. Second, LOD sensitivity may have been affected by the lack of LOB subtraction, whereas the LOQ value may have been not, maintaining its predictive value.

We are aware of the preliminary nature of our report and of its possible weaknesses. Although, the observation that an MRD estimation system independent from a pre-established threshold performs as well as in retrospective and prospective contexts is *per se* relevant, even though far from representing the identification of an absolute threshold. This proof of principle will become standard of care when its predictive value is demonstrated in different series of patients, treated with different schedules. Meanwhile, all MRD driven clinical studies should rigorously comply with the procedures recommended for rare events acquisition. In our analysis, increasing the numbers of events acquired (>500'000) and refining the population under investigation (gating CD45+ cells) resulted in a significantly enhanced predictive power of the test.

Thresholds for MRD estimation are likely to change in the near future but making them clinically informative requires that for every individual determination, the detection and quantification limits is described. Along this direction, MFC analyses in AML would possibly reach values of sensitivity comparable to those of PCR, as demonstrated in ALL and multiple myeloma.[4]

REFERENCES

1. 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.
2. Wood B, Jevremovic D, Béné MC, Yan M, Jacobs P, Litwin V. Validation of cell-based fluorescence assays: Practice guidelines from the ICSH and ICCS - Part v - performance criteria. Cytometry B Clin Cytom. 2013;84(5):315-323.
3. Wood BL. Principles of minimal residual disease detection for hematopoietic neoplasms by flow cytometry. Cytometry B Clin Cytom. 2016;90(1):47-53.
4. Theunissen P, Mejstrikova E, Sedek L, et al. Standardized flow cytometry for highly sensitive MRD measurements in B-cell acute lymphoblastic leukemia. Blood. 2017;129(3):347-357.
5. Paiva B, Puig N, Cedena MT, et al. Measurable Residual Disease by Next-Generation Flow Cytometry in Multiple Myeloma. J Clin Oncol. 2020;38(8):784-792.
6. Flores-Montero J, Sanoja-Flores L, Paiva B, et al. Next Generation Flow for highly sensitive and standardized detection of minimal residual disease in multiple myeloma. Leukemia. 2017;31(10):2094-2103.
7. Rawstron AC, Gregory WM, Tute RM De, et al. Minimal residual disease in myeloma by fl ow cytometry : independent prediction of survival bene fi t per log reduction. Blood. 2015;125(12):1932-1936.
8. Rawstron AC, Fazi C, Agathangelidis A, et al. A complementary role of multiparameter flow cytometry and high-throughput sequencing for minimal residual disease detection in chronic lymphocytic leukemia: An European Research Initiative on CLL study. Leukemia. 2016;30(4):929-936.
9. Wood BL. Acute Myeloid Leukemia Minimal Residual Disease Detection: The Difference from Normal Approach. Curr Protoc Cytom. 2020;93(1):e73.
10. Schuurhuis GJ, Heuser M, Freeman S, et al. Minimal/measurable residual disease in AML: consensus document from ELN MRD Working Party. Blood. 2018;131(12):1275-1291.
11. Hanekamp D, Bachas C, van de Loosdrecht A, Ossenkoppele G, Cloos J. Re: Myeloblasts in normal bone marrows expressing leukaemia-associated immunophenotypes. Pathology. 2020;52(2):289-291.
12. Venditti A, Buccisano F, Poeta G Del, et al. Level of minimal residual disease after consolidation therapy predicts outcome in acute myeloid leukemia : Presented in part at the 41st Annual Level of minimal residual disease after consolidation therapy predicts outcome in acute myeloid leukemia. Blood. 2000;96(12):3948-3952.
13. Buccisano F, Maurillo L, Gattei V, et al. The kinetics of reduction of minimal residual disease impacts on duration of response and survival of patients with acute myeloid leukemia. Leukemia. 2006;20(10):1783-1789.
14. 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.
15. 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.
16. 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.
17. 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.
18. 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.
19. Maurillo L, Buccisano F, Piciocchi A, et al. Minimal residual disease as biomarker for optimal biologic dosing of ARA-C in patients with acute myeloid leukemia. Am J Hematol. 2015;90(2):125-131.
20. Buccisano F, Maurillo L, Gattei V, et al. The kinetics of reduction of minimal residual disease impacts on duration of response and survival of patients with acute myeloid leukemia. Leukemia. 2006;20(10):1783.
21. 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.
22. Mandrekar SJ, Sargent DJ. Clinical trial designs for predictive biomarker validation: Theoretical considerations and practical challenges. J Clin Oncology. 2009;27(24):4027-4034.
23. Buccisano F, Maurillo L, Del Principe MI, et al. Prognostic and therapeutic implications of minimal residual disease detection in acute myeloid leukemia. Blood. 2012;119(2):332-341.
24. Arroz M, Came N, Lin P, et al. Consensus guidelines on plasma cell myeloma minimal residual disease analysis and reporting. Cytom B Clin Cytom. 2016;90(1):31-39.
25. 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.
26. Hedley BD, Keeney M. Technical issues: Flow cytometry and rare event analysis. Int J Lab Hematol. 2013;35(3):344-350.
27. Rawstron AC, Böttcher S, Letestu R, et al. Improving efficiency and sensitivity: European research initiative in CLL (ERIC) update on the international harmonised approach for flow cytometric residual disease monitoring in CLL. Leukemia. 2013;27(1):142-149.
28. Othus M, Gale RP, Hourigan CS, Walter RB. Statistics and measurable residual disease (MRD) testing: uses and abuses in hematopoietic cell transplantation. Bone Marrow Transplant. 2020;55(5):843-850.
29. 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.
30. Ossenkoppele G, Schuurhuis GJ. MRD in AML: does it already guide therapy decision-making? Hematol Am Soc Hematol Educ Progr. 2016;2016(1):356-365.
31. Hematologic Malignancies: Regulatory Considerations for Use of Minimal Residual Disease in Development of Drug and Biological Products for Treatment | FDA. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/hematologic-malignancies-regulatory-considerations-use-minimal-residual-disease-development-drug-and (accessed May 2, 2020).
32. Hourigan CS, Gale RP, Gormley NJ, Ossenkoppele GJ, Walter RB. Measurable residual disease testing in acute myeloid leukaemia. Leukemia. 2017;31(7):1482-1490.
33. Zalcberg I, D'Andrea MG, Monteiro L, Pimenta G, Xisto B. Multidisciplinary diagnostics of chronic lymphocytic leukemia: European Research Initiative on CLL - ERIC recommendations. Hematol Transfus Cell Ther. 2020;42(3):269-274.

34. 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 [Epub ahead of print].
35. Hanekamp D, Tettero JM, Ossenkoppele GJ, et al. AML/Normal Progenitor Balance Instead of Total Tumor Load (MRD) Accounts for Prognostic Impact of Flowcytometric Residual Disease in AML. Cancers (Basel). 2021;13(11):2597.

Table 1. General characteristics of the study population

	Level	Overall
Number		261
Sex (%)	Male	139 (53.3)
	Female	122 (46.7)
Age (median years [range])		49.39 [18.32, 60.95]
White blood cells (median [range])		12.66 [0.16, 186.00]
Platelets (median [range])		55.00 [7.00, 1020.00]
Risk category (%)*	NCCN-FR	87 (33.3)
	NCCN-IR	77 (29.5)
	NCCN-PR	97 (37.2)
Cytogenetic Risk (%)**	Favorable risk	28 (12.3)
	Poor risk	29 (12.8)
	Intermediate risk	170 (74.9)
FLT3 ITD (%)	Negative	190 (73.1)
	Positive	70 (26.9)
NPM1 (%)	Negative	145 (55.6)
	Positive	115 (44.1)
Graft, Number (%)	No graft	85 (32.6)
	ASCT	93 (35.6)
	AuSCT	83 (31.8)

Abbreviations: ASCT, allogeneic stem cell transplant; AuSCT, autologous stem cell transplant

*Genetic/cytogenetic risk group was attributed according to National Comprehensive Cancer Network clinical practice guidelines (version 2009) as follows: "favorable" risk [cases with Inv(16), t(8;21), t(16;16), RUNX1/RUNXT1 without c-Kit mutations, CBFB/MYH11 without c-Kit mutations, NPM1 mutation without FLT3 mutations]; "intermediate" risk [cases with Normal karyotype, isolated +8, isolated t(9;11), other karyotypic abnormalities not listed as favorable or adverse, RUNX1/RUNXT1 with c-Kit mutations, CBFB/MYH11 with c-Kit mutations, no NPM1 mutations, No FLT3-ITD mutations]; "adverse" risk [cases with complete karyotype e.g. >3 abnormalities, -5/5q-, -7/7q-, abnormalities of 11q23 excluding t(9;11), inv(3), t(3;3), t(6;9), FLT3-ITD mutations]

** Patients were stratified according to Refined Medical Research Council (MRC) classification of cytogenetic risk, as follows: "favorable"risk [cases with t(8;21), t(15;17) or inv(16)/t(16;16)]; "adverse"risk [cases with complex cytogenetic changes (> 3 unrelated abnormalities), -5, add(5q)/del(5q), -7/add(7q), t(6;11), t(10;11), t(9;22), -17, abn(17p) with other changes, 3q abnormalities excluding t(3;5), inv(3)/t(3;3)]; and "intermediate"risk [cases with normal karyotype and other non-complex];

Table 2. Integration of the "relative" 0.035% and "absolute" LOD/LOQ approaches for MRD determination

	Level	$MRD^{<0.035\%}$ LOD^{neg}/LOD^{pos}-LOQ^{neg}	$MRD^{<0.035\%}$ LOQ^{pos}	$MRD^{\geq 0.035\%}$ LOQ^{pos}	P
Number		115	39	105	
Sex (%)	Male	61 (53.0)	19 (48.7)	57 (54.3)	0.837
	Female	54 (47.0)	20 (51.3)	48 (45.7)	
Age (median [range])		48.7 [18.3, 60.3]	44.5 [21.9, 60.7]	52.3 [19.4, 60.9]	0.066
Wbc (median [range])		9.60 [0.16, 181.38]	11.70 [0.74, 186.00]	16.73 [0.48, 158.30]	0.078
Category risk (%)*	NCCN-FR	42 (36.5)	7 (17.9)	38 (36.2)	<0.001
	NCCN-IR	32 (27.8)	24 (61.5)	20 (19.0)	
	NCCN-PR	41 (35.7)	8 (20.5)	47 (44.8)	
Citogenetic risk (%)**	Favorable-risk	19 (18.8)	2 (6.7)	7 (7.4)	0.141
	Poor-risk	11 (10.9)	4 (13.3)	13 (13.8)	
	Intermediate-risk	71 (70.3)	24 (80.0)	74 (78.7)	
FLT3-ITD (%)	Negative	83 (72.8)	35 (89.7)	71 (67.6)	0.028
	Positive	31 (27.2)	4 (10.3)	34 (32.4)	
NPM1 (%)	Negative	64 (55.7)	30 (76.9)	50 (48.1)	0.008
	Positive	51 (44.3)	9 (23.1)	54 (51.9)	
Graft, Number (%)	No graft	38 (33.0)	9 (23.1)	37 (35.2)	0.629
	ASCT	38 (33.0)	17 (43.6)	37 (35.2)	
	AuSCT	39 (33.9)	13 (33.3)	31 (29.5)	

Abbreviations: LOD, limit of detection; LOQ, limit of quantification; MRD, measurable residual disease; Wbc, white blood cells; NCCN, National Comprehensive Cancer Network; ASCT, allogeneic stem cell transplant; AuSCT, autologous stem cell transplant

*Genetic/cytogenetic risk group was attributed according to National Comprehensive Cancer Network clinical practice guidelines (version 2009) as follows: "favorable" risk [cases with Inv(16), t(8;21), t(16;16), RUNX1/RUNXT1 without c-Kit mutations, CBFB/MYH11 without c-Kit mutations, NPM1 mutation without FLT3 mutations]; "intermediate" risk [cases with Normal karyotype, isolated +8, isolated t(9;11), other karyotypic abnormalities not listed as favorable or adverse, RUNX1/RUNXT1 with c-Kit mutations, CBFB/MYH11 with c-Kit mutations, no NPM1 mutations, No FLT3-ITD mutations]; "adverse" risk [cases with complete karyotype e.g. >3 abnormalities, -5/5q-, -7/7q-, abnormalities of 11q23 excluding t(9;11), inv(3), t(3;3), t(6;9), FLT3-ITD mutations]

** Patients were stratified according to Refined Medical Research Council (MRC) classification of cytogenetic risk, as follows: "favorable" risk [cases with t(8;21), t(15;17) or inv(16)/t(16;16)]; "adverse" risk [cases with complex cytogenetic changes (> 3 unrelated abnormalities), -5, add(5q)/del(5q), -7/add(7q), t(6;11), t(10;11), t(9;22), -17, abn(17p) with other changes, 3q abnormalities excluding t(3;5), inv(3)/t(3;3)]; and "intermediate" risk [cases with normal karyotype and other non-complex];

Table 3. Univariate and multivariate cox regression model for OS

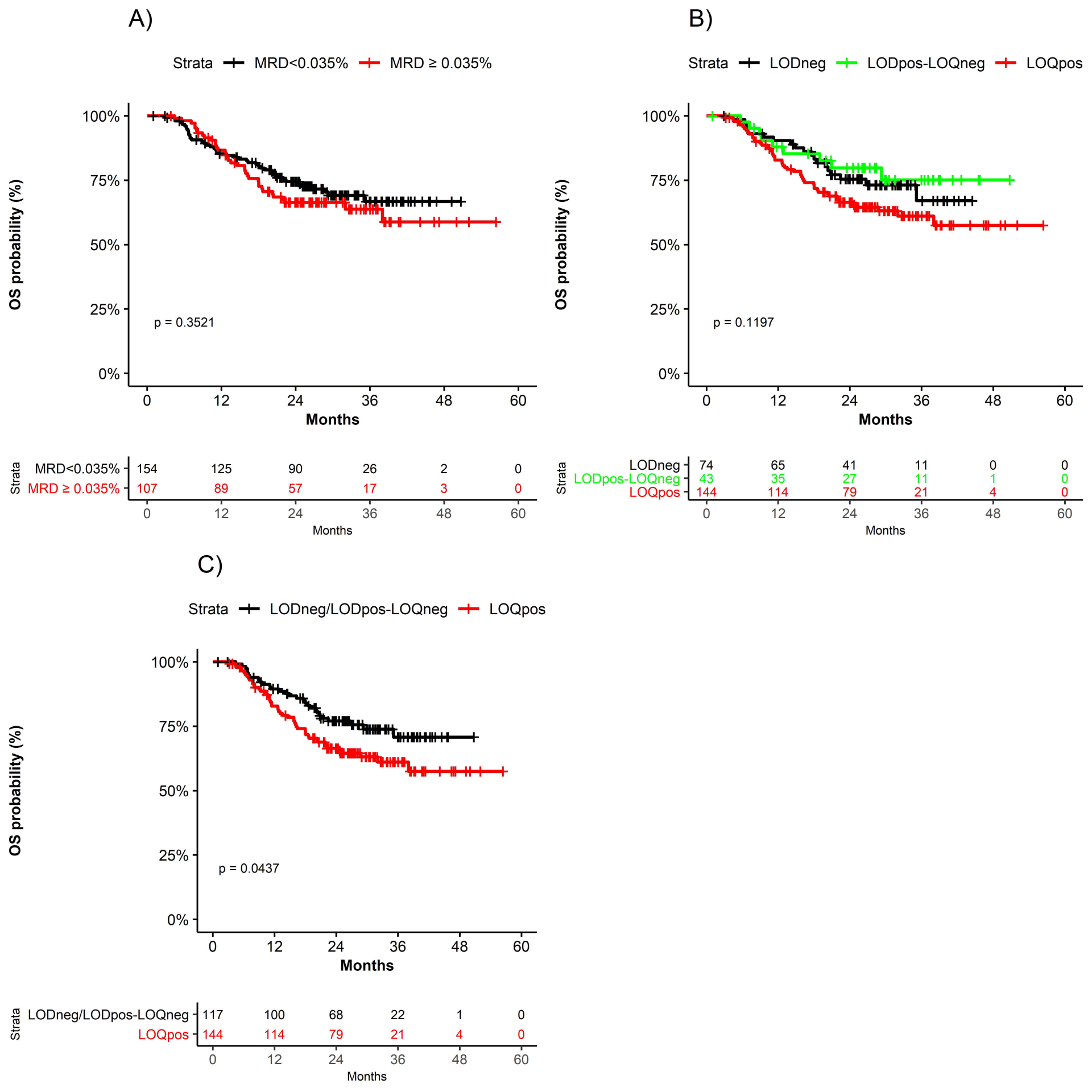
	Univariate			Multivariate		
Characteristic	*HR*	*95% CI*	*p-value*	*HR*	*95% CI*	*p-value*
WBC	1.00	0.99, 1.00	0.30			
FLT3 ITD						
Negative	—	—				
Positive	2.40	1.53, 3.77	**<0.001**			
Risk category						
NCCN-FR	—	—		—	—	
NCCN-IR	1.95	1.00, 3.81	**0.051**	1.99	0.95, 4.15	0.068
NCCN-PR	3.73	2.04, 6.82	**<0.001**	5.02	2.31, 10.9	**<0.001**
LOD LOQ stratification						
LOD^{neg}-$LOD^{pos}LOQ^{neg}$	—	—				
LOQ^{pos}	1.60	1.01, 2.54	**0.046**			
BM MRD STATUS POST CONS						
Negative	—	—				
Positive	1.23	0.79, 1.92	0.35			
MRD_LODLOQ						
$MRD^{<0.035\%}$/ LOD^{neg}-$LOD^{pos}LOQ^{neg}$	—	—		—	—	
$MRD^{\geq 0.035\%}$/ LOD^{neg}-$LOD^{pos}LOQ^{neg}$	0.00	0.00, Inf	>0.99	0.00	0.00, Inf	>0.99
$MRD^{<0.035\%}$/ LOQ^{pos}	1.82	0.97, 3.41	**0.061**	2.19	1.13, 4.27	**0.021**
$MRD^{\geq 0.035\%}$/ LOQ^{pos}	1.49	0.91, 2.44	0.11	1.29	0.78, 2.13	0.33
graftYN						
No graft	—	—		—	—	
allo	0.77	0.47, 1.27	0.30	0.47	0.28, 0.80	**0.005**
auto	0.41	0.23, 0.75	**0.003**	0.72	0.36, 1.46	0.37

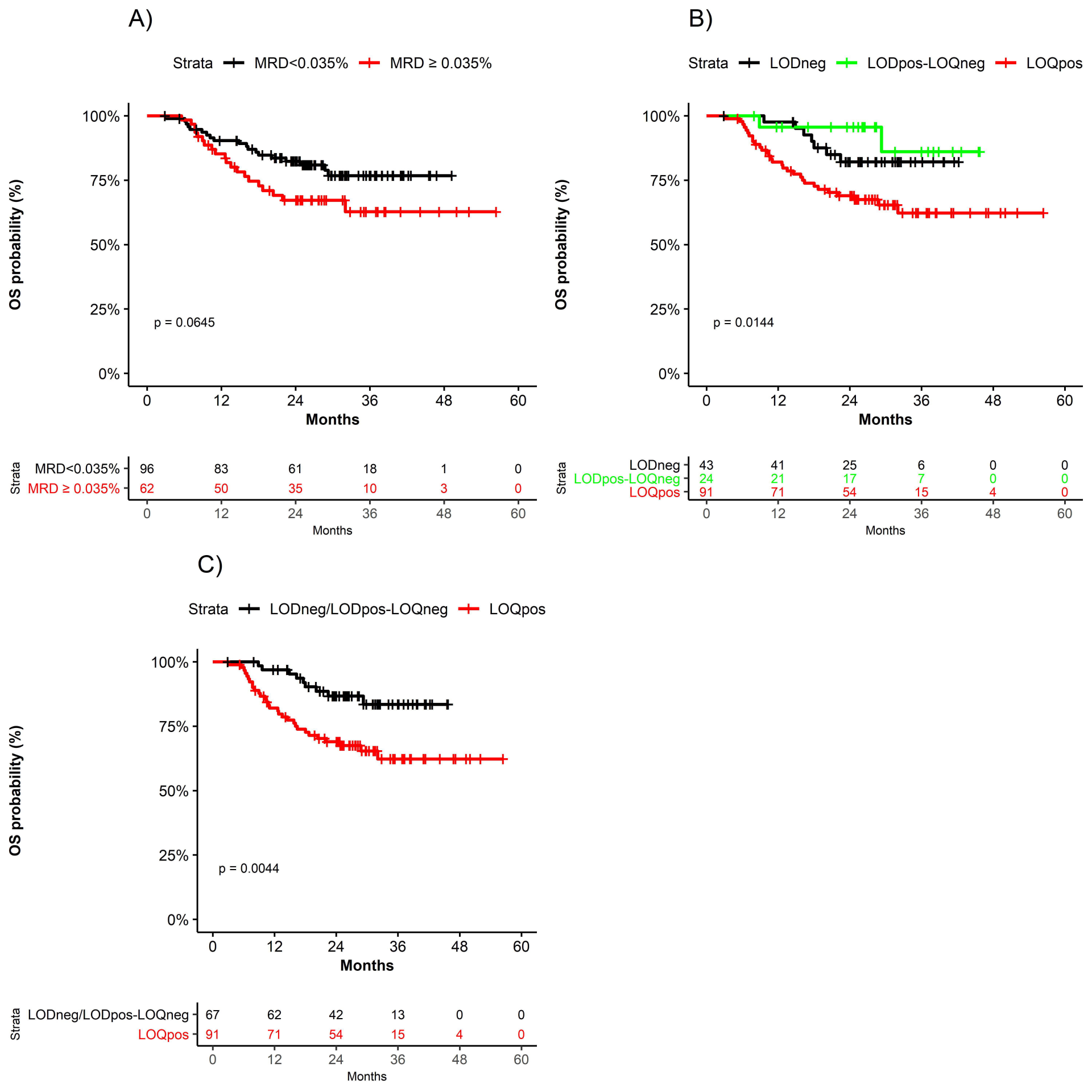
Abbreviations: LOD, limit of detection; LOQ, limit of quantification; MRD, measurable residual disease; Wbc, white blood cells; NCCN, National Comprehensive Cancer Network; ASCT, allogeneic stem cell transplant; AuSCT, autologous stem cell transplant; HR = Hazard Ratio, CI = Confidence Interval

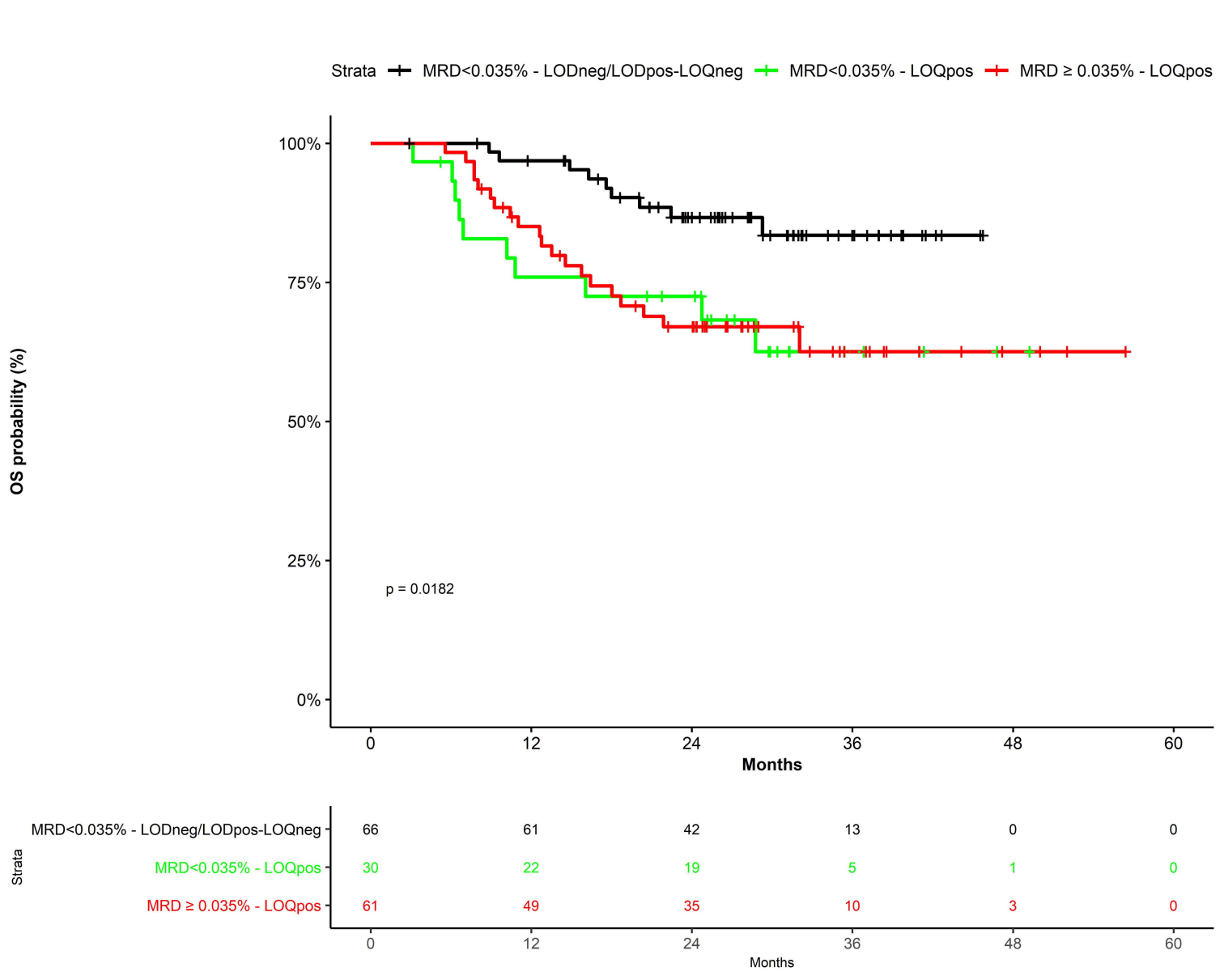
Figure 1. Overall Survival analysis of the whole series of 261 patients according to different MRD estimates. MRD stratification according to the AML1310 threshold (0.035%) was not statistically different (**A**). LOD^{neg}, LOD^{pos}-LOQ^{neg} and LOQ^{pos} are analyzed separately (**B**) and merging LOD^{neg} and LOD^{pos}-LOQ^{neg} (**C**), with only the latter reaching a statistically significant difference (p=0.043).

Figure 2. Overall Survival analysis of 158 patients where >500'000 CD45+ have been acquired. Stratification according to the AM1310 MRD threshold showed a lower power of discrimination in terms of 2-year OS (p=0.064) (**A**). LOD^{neg}, LOD^{pos}-LOQ^{neg} and LOQ^{pos} are analyzed separately (**B**) and merging LOD^{neg} and LOD^{pos}-LOQ^{neg} (**C**), both tests stratified patients with a statistical significance (p=0.023 and p=0.009, respectively).

Figure 3. Overall Survival analysis of $MRD^{<0.035\%}$ and $MRD^{\geq 0.035\%}$ patients according to the LOD-LOQ status. $MRD^{<0.035\%}LOD^{neg}$/LOD^{pos}-LOQ^{neg} patients had a longer duration of OS as compared to $MRD^{<0.035\%}LOQ^{pos}$ and $MRD^{\geq 0.035\%}LOQ^{pos}$ (p=0.018). Even more, $MRD^{<0.035\%}$ patients had a statistically different OS if they tested LOD^{neg}/LOD^{pos}-$LOQ^{neg\ neg}$ or LOQ^{pos} (p=0.007).


A)
Strata MRD<0.035% MRD ≥ 0.035%
OS probability (%)
100%
75%
50%
25%
0%
p = 0.3521
0 12 24 36 48 60
Months
Strata
MRD<0.035% 154 125 90 26 2 0
MRD ≥ 0.035% 107 89 57 17 3 0
0 12 24 36 48 60
Months
B)
Strata LODneg LODpos-LOQneg LOQpos
OS probability (%)
100%
75%
50%
25%
0%
p = 0.1197
0 12 24 36 48 60
Months
Strata
LODneg 74 65 41 11 0 0
LODpos-LOQneg 43 35 27 11 1 0
LOQpos 144 114 79 21 4 0
0 12 24 36 48 60
Months
C)
Strata LODneg/LODpos-LOQneg LOQpos
OS probability (%)
100%
75%
50%
25%
0%
p = 0.0437
0 12 24 36 48 60
Months
Strata
LODneg/LODpos-LOQneg 117 100 68 22 1 0
LOQpos 144 114 79 21 4 0
0 12 24 36 48 60
Months


A)
Strata MRD<0.035% MRD ≥ 0.035%
OS probability (%)
100%
75%
50%
25%
0%
p = 0.0645
0 12 24 36 48 60
Months
Strata
MRD<0.035% 96 83 61 18 1 0
MRD ≥ 0.035% 62 50 35 10 3 0
0 12 24 36 48 60
Months
B)
Strata LODneg LODpos-LOQneg LOQpos
OS probability (%)
100%
75%
50%
25%
0%
p = 0.0144
0 12 24 36 48 60
Months
Strata
LODneg 43 41 25 6 0 0
LODpos-LOQneg 24 21 17 7 0 0
LOQpos 91 71 54 15 4 0
0 12 24 36 48 60
Months
C)
Strata LODneg/LODpos-LOQneg LOQpos
OS probability (%)
100%
75%
50%
25%
0%
p = 0.0044
0 12 24 36 48 60
Months
Strata
LODneg/LODpos-LOQneg 67 62 42 13 0 0
LOQpos 91 71 54 15 4 0
0 12 24 36 48 60
Months


Strata
MRD<0.035% - LODneg/LODpos-LOQneg
MRD<0.035% - LOQpos
MRD ≥ 0.035% - LOQpos
OS probability (%)
100%
75%
50%
25%
0%
p = 0.0182
0
12
24
36
48
60
Months
MRD<0.035% - LODneg/LODpos-LOQneg 66 61 42 13 0 0
MRD<0.035% - LOQpos 30 22 19 5 1 0
MRD ≥ 0.035% - LOQpos 61 49 35 10 3 0
Strata
0
12
24
36
48
60
Months

Supplemental methods

AML1310 study design and MRD assessment.

Patients with newly-diagnose AML were eligible for the GIMEMA AML1310 trial provided they met the following criteria for eligibility: age 18 to 60.9 years; (ii) AML other than M3; (iii) WHO performance status 0-3; (iv) adequate liver (serum bilirubin level ≤ 2 UNL; AST and ALT ≤ 3 UNL) and renal (serum creatinine ≤ 2 UNL) functions; (v) LVEF ≥ 50% by echocardiogram; (vi) absence of severe concomitant neurological or psychiatric diseases and congestive heart failure or active uncontrolled infections; (vii) signed informed consent. Patients with therapy-related AML were not considered eligible. Exclusion criteria included blast crisis of chronic myeloid leukemia, AML supervening after other chronic myeloproliferative diseases or antecedent myelodysplastic syndromes of more than six months duration and other progressive malignant diseases.

Study procedures included upfront evaluation included bone marrow (BM) aspirate for morphology, cytogenetics, molecular genetics and MFC analysis. The baseline MFC assessment was a necessary step, not only for diagnostic purposes, but also to identify leukemia associated immunophenotypes (LAIP). Identification of baseline LAIPs was the essential requirement for monitoring MRD after therapy. Patients were studied at diagnosis for the presence of RUNX1-RUNX1T1 or CBFβ/MYH11 rearrangements, defining core binding factor (CBF) leukemias, and for NPM1, FLT3 and c-KIT mutations. Molecular analysis, LAIPs assessment and post-consolidation MRD determinations were centralized at Laboratorio di Diagnostica Integrata Oncoematologica “OPPO”, at Tor Vergata University Hospital of Rome, whereas conventional karyotype was carried out at local institutions. Response to treatment was assessed on BM and peripheral blood, according to the recommendations of an international working group. Patients who did not achieve complete remission (CR), CR incomplete (CRi) or partial remission (PR) after the first induction course or CR/CRi after two induction courses were considered as treatment failures. At the established time-point, BM MRD was determined by a high-sensitivity 8–color MFC assay. BM samples were processed following the Lyse-wash-stain-wash procedure. This includes bulk lysis of red blood cells followed by washing with PBS, resuspension of the pellet in a smaller volume allowing for increased cell concentration and staining of the cells with MoAb cocktail. The bulk lysis leads to a higher reproducibility since the labelling conditions are reproducible and the volume is constant for a given quantity of cells. Lysis was performed by Ammonium chloride (NH4Cl) because of its minimal effects on cell physical properties. The threshold for discriminating MRD negative from MRD positive cases was set at 3.5x10-4 (0.035%) RLC and the selected time-point was the post-consolidation phase,

once the hematologic recovery was complete. MRD assessment was carried on the total number of mononuclear cells (MNC). The AML1310 trial was designed at a time when ELN 2010/2017 and NCCN 2018 recommendations were not yet published. Therefore, when the trial regulatory path was concluded, we started recruiting and stratifying patients according to contemporary classification, that was the NCCN 2009 version 1.27 For the purpose of our AML1310 study, 4 categories of risk were identified: 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. Moreover, we enucleated a fifth group of patients belonging to the intermediate-risk category, in whom we failed to identify any LAIP (NCCN-IR-no-LAIP category); these patients were allocated to the AuSCT post-consolidation option. ASCT and AuSCT were to be performed within three months of the end of the consolidation course.

Table 1S. Total number of events acquired according to LOD/LOQ coding and details on MNC and CD45 expressing cells denominators

	Level	Overall
Number. of MNC acquired (median [range])		559'197 [100'450-1'561'221]
Number of CD45 expressing cells acquired (median [range])		538'527 [88'040-1'548'172]
Number of positive events (median [range])		89 [0-38'061]
Percentage of positive events (median [range])		0.0161 [0.0000-9.0152]
LOD on CD45 expressing cells (median [range])		0.0037 [0.0013-0.0227]
LOQ on CD45 expressing cells (median [range])		0.0093 [0.0032-0.0568]
Positivity according to LOD LOQ on CD45 expressing cells (%)	LOD^{neg}	74 (28.4)
	$LOD^{pos}LOQ^{neg}$	43 (16.5)
	LOQ^{pos}	144 (55.1)

Abbreviations: LOD, limit of detection; LOQ, limit of quantification; MNC, mononuclear cells.

Table 2S. Interactions between the "relative" 0.035% and "absolute" LOD/LOQ MRD estimates

Level	MRD Negative*	MRD Positive**	P
CD45-Positive LOD^{neg}	74 (48.1%)	0 (0.0%)	<0.001
CD45-Positive $LOD^{pos}LOQ^{neg}$	41 (26.6%)	2 (1.9%)	
CD45-Positive LOQ^{pos}	39 (25.3%)	105 (98.1%)	

Abbreviations: LOD, limit of detection; LOQ, limit of quantification; MRD, measurable residual disease.
*Residual leukemic cells <0.035%
**Residual leukemic cells >0.035%

Figure 1S. Overall Survival analysis of MRD<0.035% and MRD≥0.035% patients according to the LOD-LOQ status in the NCCN intermediate risk group. MRD<0.035%LOD^{neg}/LOD^{pos}LOQ^{neg} patients had a longer duration of OS as compared to MRD<0.035%LOQ^{pos} and MRD≥0.035%LOQ^{pos} (p=0.0146).

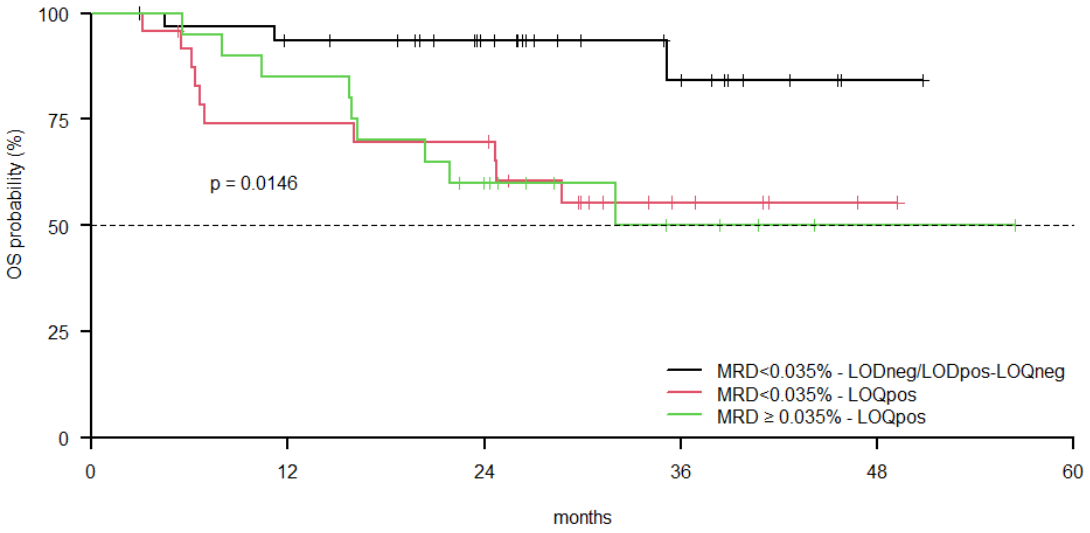


Figure 2S. Interaction of the 3 LOD-LOQ categories with post-remissional treatment (AuSCT, ASCT and no graft). LOD^{neg}/LOD^{pos}-LOQ^{neg} patients submitted to AuSCT had the best 2-years OS (88.9%) as compared to all the other possible combinations of LOD-LOQ and treatment (p=0.026).

A