haematologica
Journal of The Ferrata Storti Foundation

# Effect of age and treatment on predictive value of measurable residual disease: implications for clinical management of adult patients with acute myeloid leukemia

by Francesco Mannelli, Matteo Piccini, Sara Bencini, Giacomo Gianfaldoni, Benedetta Peruzzi, Roberto Caporale, Barbara Scappini, Laura Fasano, Elisa Quinti, Gaia Ciolli, Andrea Pasquini, Francesca Crupi, Sofia Pilerci, Fabiana Pancani, Leonardo Signori, Danilo Tarantino, Chiara Maccari, Vivian Paradiso, Francesco Annunziato, Paola Guglielmelli, and Alessandro Maria Vannucchi

*Received: March 24, 2023.*
*Accepted: June 15, 2023.*

*Citation: Francesco Mannelli, Matteo Piccini, Sara Bencini, Giacomo Gianfaldoni, Benedetta Peruzzi, Roberto Caporale, Barbara Scappini, Laura Fasano, Elisa Quinti, Gaia Ciolli, Andrea Pasquini, Francesca Crupi, Sofia Pilerci, Fabiana Pancani, Leonardo Signori, Danilo Tarantino, Chiara Maccari, Vivian Paradiso, Francesco Annunziato, Paola Guglielmelli, and Alessandro Maria Vannucchi. Effect of age and treatment on predictive value of measurable residual disease: implications for clinical management of adult patients with acute myeloid leukemia. Haematologica. 2023 June 22. doi: 10.3324/haematol.2023.283196 [Epub ahead of print]*

# Effect of age and treatment on predictive value of measurable residual disease: implications for clinical management of adult patients with acute myeloid leukemia

Francesco Mannelli[1,2], Matteo Piccini[1], Sara Bencini[3], Giacomo Gianfaldoni[1], Benedetta Peruzzi[3], Roberto Caporale[3], Barbara Scappini[1], Laura Fasano[1], Elisa Quinti[1], Gaia Ciolli[1], Andrea Pasquini[1], Francesca Crupi[1], Sofia Pilerci[1], Fabiana Pancani[1,2], Leonardo Signori[1,2], Danilo Tarantino[1,2], Chiara Maccari[1,2], Vivian Paradiso[1], Francesco Annunziato[3], Paola Guglielmelli[1,2], Alessandro M. Vannucchi[1,2]

[1]SOD Ematologia, Università di Firenze, AOU Careggi, Firenze, Italy; [2]Centro Ricerca e Innovazione Malattie Mieloproliferative (CRIMM), AOU Careggi, Firenze, Italy; [3]Centro Diagnostico di Citofluorimetria e Immunoterapia, AOU Careggi, Firenze, Italy

**Corresponding author:** Francesco Mannelli, MD, SOD Ematologia, Centro Ricerca e Innovazione Malattie Mieloproliferative (CRIMM), Università di Firenze, AOU Careggi, Largo Brambilla 3, 50134 Firenze, Italy; Phone/fax: +39 055 7947824; e-mail: francesco.mannelli@unifi.it

**Keywords**: acute myeloid leukemia; measurable residual disease; MRD; immunophenotype; molecular genetics; induction therapy, high dose cytarabine

Text word count: 4080
Abstract word count: 246
No. of figures/tables: 4/2
No. of references: 38

**Acknowledgments**

This work is dedicated to the memory of Dr. Franco Leoni, for his invaluable mentorship in the care of patients with acute leukemias.

The work was supported by AIRC 5x1000 call "Metastatic disease: the key unmet need in oncology" to MYNERVA project, #21267 (MYeloid NEoplasms Research Venture AIRC) and from Legato Zottola. A detailed description of the MYNERVA project is available at http://www.progettomynerva.it

**Authorship Contributions**

FM designed and performed the research, analyzed, and interpreted the data, and wrote the manuscript. MP performed the research, analyzed, and interpreted the data. SB performed flow cytometric analysis and collected the data. BP, RC performed flow cytometric analysis. GG, BS, LS, EQ, GC, FC, SP, VP collected clinical data. FP, LS, CM, PG performed molecular analysis. FA and AMV interpreted the data and critically reviewed the manuscript.

**Conflicts of Interest Disclosures**

The authors declare no competing financial interests.

**Data sharing**

The data that support the findings of this study are available on request from the corresponding author.

**Abstract**

Measurable residual disease (MRD) is a powerful predictor of outcome in acute myeloid leukemia (AML). In the early phases of treatment, MRD refines initial disease risk stratification and is used for the allocation to allogeneic transplant (HSCT). Despite its well-established role, a relatively high fraction of patients eventually relapses albeit achieving $MRD_{neg}$ status. The aim of this work was to assess specifically the influence of baseline features and treatment intensity on the predictive value of an $MRD_{neg}$ status, particularly focusing on MRD2, measured after 2 consecutive chemotherapeutic cycles. Among baseline features, younger $MRD2_{neg}$ patients (<55 y) had a significantly longer DFS (median not reached) compared to their elderly counterpart (median 25.0 months, P=0.013, HR=2.08). Treatment intensity, specifically the delivery of high dose of ARA-C in induction or first consolidation, had apparently a pejorative effect on the outcome of $MRD2_{neg}$ patients compared to standard dose (P=0.048, HR=1.80), a finding confirmed also by the analysis of data extracted from the literature. The combination of age and treatment intensity allowed us to identify categories of patients, among those who reached a $MRD2_{neg}$ status, characterized by significantly different disease-free survival rate. Our data showed that variables such as age and intensity of treatment administered can impact on the predictive value of MRD in patients with AML. In addition to underscoring the need for further improvement of MRD analysis, these findings call for a reasoned application of MRD data, as currently available, to modulate consolidation therapy on adequately estimated relapse rates.

## Introduction

The determination of measurable residual disease (MRD) relies on the detection of leukemic cells below the morphology-based threshold that establishes complete remission definition (CR)[1]. The persistence of MRD is a powerful predictor of disease relapse and poor outcome in acute myeloid leukemia (AML)[2–6], a role formally recognized in the European Leukemia Net (ELN) recommendations, which defines a specific response category of CR without MRD ($CR_{MRD-}$)[1]. The prognostic impact of MRD measurement in AML has been demonstrated regardless of the method applied[6], although the ELN consensus suggests employing a molecular probe for core binding factor (CBF) related and *NPM1*-mutated AML, and a multi-parameter flow cytometry (MFC) approach in all other subgroups[7].

In the early phases of treatment, MRD assessment refines initial disease risk stratification and is primarily used for the allocation of eligible patients to allogeneic transplant (HSCT), generally reserved for categories deemed to be at high risk of relapse[8]. From this perspective, the most informative time-point is set after two cycles of intensive chemotherapy (i.e., MRD2), as it has been shown to yield the greatest predictive value on outcome[3,5,9]. Such a strategy is stated as standard by ELN guidelines[7,10] and is increasingly becoming the basis for clinical decision-making in AML management[9,11,12].

Despite its indisputable role, the prognostic value of MRD may be influenced by the clinical context, including age range[13] and the genetic profile[14], as well as by pre-MRD intensity of treatment. In fact, the predictive value of MRD2 reflects the degree of chemosensitivity of leukemic cells: its positioning after two chemotherapy cycles provides a reliable estimation of the likelihood that further chemotherapy may lead to disease eradication. Overall, the reliability of MRD2 largely depends on previously delivered treatment, including the dosages of the backbone agents (cytarabine and anthracycline), the use of a third drug alongside (etoposide, fludarabine, gemtuzumab ozogamicin and, more recently for *FLT3*-mutant cases, an FLT3 inhibitor), as well as the regimen schedule (standard, double induction, sequential). Therefore, MRD2 status somehow captures the information on disease sensitivity to the treatment actually delivered. Furthermore, MRD is measured after achievement of CR following induction therapy, and the intensity of chemotherapy in the first cycle is known to influence the rate of CR[15–17]. In other words, different treatment intensity in the first cycle contributes to shape the population of patients in which MRD is later assessed starting from CR achievement onward, and this might affect the ability of MRD itself to estimate outcome. Although the influence of such effect is systematically stated in guidelines[7], its impact has not been fully evaluated and is generally not taken into account in clinical decision-making. The aim of this work was to assess specifically the influence of baseline features (age,

white blood cell count, ELN stratification), and treatment intensity on the predictive value of MRD, specifically focusing on MRD2 negative status owing to the key clinical information it can provide. For this purpose, we retrospectively analyzed our database and then interrogated the available literature to validate our findings.

## Methods

*Patients*

Patients entering the study had a diagnosis of non-promyelocytic AML, based on morphological, immunophenotypic, and molecular criteria. Once eligible for intensive chemotherapy, they were consecutively assigned to treatment based on the availability of a clinical trial at the time of diagnosis and based on local institutional treatment choices as previously reported[18] and detailed in the Supplemental Data. For the purpose of the study, the first two chemotherapy cycles were categorized for treatment intensity according to the dosage of cytarabine (ARA-C): standard dose (SDAC) included "3+7" and "3+7"-like regimens sharing a single dose infusion of 100-200 mg/sqm and a cumulative dose of up to 1.400 mg/sqm per cycle; high dose (HDAC) included several schedules sharing a single dose of at least 1.000 mg/sqm and a cumulative dose of at least 6.000 mg/sqm (Supplemental Data). CPX-351 was categorized as SDAC.

The study was approved by the local institutional review board (protocol number: 2013/0021560). Written informed consent was obtained from study patients in accordance with the Declaration of Helsinki. Enrolment criteria required (1) intensive treatment, (2) CR achievement after first induction therapy, (3) availability of MRD data after induction (MRD1) and first consolidation cycle (MRD2).

*MRD study*

Multi-parameter Flow Cytometry

MFC study files reporting individual leukemia-aberrant immune-phenotype (LAIP) profiles were acquired locally according to pre-defined standard operating procedures. LAIPs were defined by antigen expression on blasts from fresh diagnostic BM (or PB in case of *punctio sicca*) samples. Standard MFC methodologies for LAIP definition are detailed in the Supplemental file and reported elsewhere[19]. MRD was expressed as the percentage of LAIP+ cells on CD45+ cells. MRD was defined as positive for any detection ≥0.1%, in accordance with ELN recommendations[7,20].

PCR-based MRD

Sensitive Real-time quantitative-polymerase chain reaction assays (RQ-PCR) was used for detection of MRD in patients with a suitable molecular probe (*RUNX1T1* and *CBFB-MYH11* transcripts and *NPM1* gene mutations). RQ-PCR was performed following the Europe Against Cancer (EAC) program recommendations[21] with a sensitivity of $10^{-5}$. MRD was defined as positive for any detection ≥0.01%. Standard PCR methodologies are detailed in the Supplemental file.

*Definitions*

CR, disease-free survival (DFS), and overall survival (OS) were defined according to standard criteria[1]. As regards prognostic stratification, the criteria used for the therapeutic decision-making across different time periods, and particularly allocation to HSCT, are specified in the Supplemental file. In post hoc analysis, patients were stratified according to the availability of genetic data according to Medical Research Council (MRC) criteria[22], and European Leukemia Net (ELN) version 2010-2017[1,23].

*Analysis of literature*

We carried out a search in PubMed for articles published between 2000 and 2021 by filtering for keywords (AML, acute myeloid leukemia, or acute myelogenous leukemia, and MRD, minimal residual disease, or measurable residual disease). The analytical method, the criteria for paper selection and the extracted data are detailed in Supplemental file (Method S3, Table S1). Based on this assessment, a total of 33 articles were selected. Then, we extracted survival data from Kaplan-Meier curves by using the commercial graph digitizer software (Digitizelt, version 2.1; Bormisoft) and applying a previously published algorithm to reconstruct survival data for $MRD_{pos}$ and $MRD_{neg}$ cases[24]. The main characteristics of treatment in the first two cycles (drugs, ARA-C dosage, schedule) were obtained for each report and tabulated as in Supplemental Table S2.

*Statistical methods*

All statistical analyses were performed using R version 3.5.0 and SPPS version 26. Statistical methods are detailed in Supplemental file (S5).

**Results**

Characteristics of patients and treatment flow

From April 2004 to January 2022, 194 patients affected by AML met the inclusion criteria. Considering the large enrolment period, we carried out an analysis of OS to check for an influence by the year of diagnosis (Figure S1): no relevant impact on survival emerged. The clinical and

biological characteristics of our patient series are detailed in Table 1. In addition to baseline features, we classified patients according to treatment intensity (SDAC-based and HDAC-containing regimens) and MRD status after induction (MRD1) and first consolidation (MRD2); patients receiving HDAC in at least one of the first two cycles, in induction (cycle 1) or consolidation (cycle 2) (n=124, 63.9%), were compared with those receiving SDAC in both cycles (n=70, 36.1%) (Figure 1).

MRD study

For CBF-related and *NPM1*-mutated AML, the categorization according to MRD was based on PCR data, available in all 33 CBF-related cases, and in 62/80 (77.5%) *NPM1*-mutated ones. In all other cases, MFC was used. After induction, 110 (56.7%) patients were defined as $MRD1_{neg}$ and 84 (43.3%) as $MRD1_{pos}$. After first consolidation, a total of 121 (62.4%) cases reached MRD2 negative status, whereas 73 (37.6%) patients resulted $MRD2_{pos}$ (Figure 1). MRD1 and MRD2 status showed impact on DFS (Figure S2, S3), with MRD2 being more effective in discriminating outcome categories (C-index 0.61 versus 0.547, P = 0.009). MRD2 status also significantly influenced OS (P=0.0014, HR=2.08) (Figure S3).

Analysis according to baseline features

Due to its greater informativeness for prognosis, further analyses were focused on MRD2. We searched for any interaction between the prognostic impact of MRD2 and baseline features. First, patients were stratified for age: the observed median age was 55 y; by means of a ROC curve analysis using relapse as a binary endpoint, we confirmed an age threshold of 55.5 y as featured by the best performance (Youden score=0.224) as the result of the combination of a sensitivity and specificity of 0.564 and 0.66, respectively. Accordingly, we separated patients for their age < 55 y and ≥ 55 y. Also, we categorized patients for gender, white blood cell count (WBC) ($<30x10^9$/L and $\geq 30x10^9$/L), and ELN (favorable and intermediate). ELN adverse category was too limited in size (n=24, of whom 13 $MRD2_{neg}$) for such an analysis. Significant differences between $MRD2_{neg}$ and $MRD2_{pos}$ patients were found as regarded DFS and OS in age-, WBC-, and ELN-defined subgroups (Figure S4-S6) and within female patient category (Figure S6). In male patients, we observed a non-significant trend in favor of $MRD2_{neg}$ group (Figure S7). Since we observed a still remarkable rate of relapse in the $MRD2_{neg}$ group (DFS rate at 2 years, 60.2% versus 32.1% in $MRD2_{pos}$ patients, P=0.00013), we searched for variables impacting on DFS by comparing the outcome of $MRD2_{neg}$ patients in different disease categories. We found that younger (i.e., <55y) $MRD2_{neg}$ patients showed a significantly longer DFS than their older (i.e., ≥55y) counterpart (P=0.013, HR=2.08,

Figure S6), while no significant differences emerged for gender- (P=0.11), WBC- (P=0.29) and ELN-(P=0.1) related strata (Figure S8).

Analysis according to treatment groups

We then analyzed the baseline characteristics and the outcome of patients stratified according to the treatment intensity (SDAC vs HDAC). The only statistically significant difference was younger age in patients treated with HDAC-containing induction (47 vs 57 y; P<0.0001) (Table S3). Also, no significant difference in outcome was observed according to cytarabine dose in induction (Figure S9).

We further stratified patients according to the cumulative effect of induction and first consolidation, i.e., those who received at least one HDAC-containing cycle versus those treated only with SDAC. We only observed a trend towards enrichment in 2010 ELN adverse risk categories within HDAC-treated group (Table 1). However, no significant outcome benefit was highlighted for patients receiving HDAC during induction or consolidation (Figure S9).

Effects of treatment intensity on MRD2 predictive capability

We searched for interactions between MRD2 status and outcome within different treatment groups. We observed statistically significant longer DFS in $MRD2_{neg}$ versus $MRD2_{pos}$ patients who received SDAC during induction and first consolidation (HR=3.87, P<0.0001, C-index=0.652, Figure 2A). At the opposite, MRD2 status failed to significantly resolve DFS in patients treated with at least one HDAC-containing cycle, in induction or consolidation (HR=1.60, P=0.066, C-index=0.585, Figure 2C). Similar findings concerned the impact of MRD2 on OS, in SDAC being associated with a significant value (P<0.0001, HR=3.85, 95%CI 1.87-7.93), unlike what was observed for HDAC (P=0.35, HR=1.33, 95%CI 0.73-2.41) (Figure 2B-D). Above differences were maintained also in the category of intermediate-risk karyotype: the favorable impact of $MRD2_{neg}$ status on DFS was more evident in patients receiving SDAC (HR=3.98, P=0.00078, C-index=0.65) than HDAC (HR=2.13, P=0.01, C-index=0.613, Figure S10), though statistical significance was reached also in the latter one. Conversely, among $MRD2_{pos}$ patients, DFS was largely unaffected by treatment arm (median 11.6 and 11.4 months in SDAC- and HDAC-treated patients, respectively) and the worse performance of MRD2 in HDAC group was largely the consequence of a relatively high rate of relapse in $MRD2_{neg}$ patients. OS estimates followed the same pattern in intermediate-risk karyotype tier (Figure S10).

Characteristics of $MRD2_{neg}$ patients according to treatment arm

We then analyzed $MRD2_{neg}$ patients according to the intensity of the first two chemotherapy cycles. At baseline, there was a statistically significant enrichment of ELN adverse-risk patients in HDAC-treated (16.9% according to 2017 version) compared to SDAC-treated $MRD2_{neg}$ cases (2.0%, P=0.014) (Table 2). When comparing patients reaching $MRD2_{neg}$ status after SDAC vs HDAC, significantly different DFS (P=0.048, HR=1.80, 95% CI 1.0-3.23) and OS (P=0.049, HR=1.94, 95% CI 0.99-3.8) were highlighted (Figure S11). After being superimposable at 6 months (93.9% and 90.0% for SDAC- and HDAC-treated patients, respectively), DFS rate started to diverge at 12 months (83.4% vs 71.5%) and were 70.6% vs 51.7% at 24 months. Similar trends in DFS were observed in HDAC vs SDAC patients with intermediate-risk karyotype (P=0.092) and ELN 2017 (P=0.23) categories (Figure S12).

Combined model including age and treatment intensity

Based on the observed influence of age and treatment on the prognostic value of $MRD2_{neg}$ status, we combined these two variables and stratified $MRD2_{neg}$ patients accordingly. Twenty-four (out of 121, 19.8%) patients were younger than 55y and reached MRD2 negativity after two SDAC-based cycles (COMB_1): they showed a DFS rate of 86.4% at 3 years. At the opposite, the group of 38 (31.4%) elderly patients who had received at least one HDAC-based cycle (COMB_3) showed 22.2 months and 46.6% as median DFS and DFS rate at 3 years, respectively (Figure 3A-B). The remainder patients (COMB_2) displayed an intermediate behavior, but with DFS largely superimposable to the latter group in the first two years after CR achievement (Figure 3A). After combining COMB_2 and _3 patients, survival estimates for DFS (P=0.0034, HR=4.13) and OS (P=0.0014, HR=7.36) were significantly different in respect to COMB-1 patients (Figure 3C-D). The significant effect of the model on DFS was maintained in intermediate-risk stratification tiers according to karyotype (P=0.013, Figure S13) and ELN 2017 (P=0.016, Figure S13).

Allogeneic transplant

Further analysis included the censoring for patients who received HSCT. The differential effect of age and treatment intensity on DFS of $MRD2_{neg}$ patients was confirmed in this further analysis, either as single variables (Figure S14) or in the combined model (Figure S15). A non-significant trend for a lower actual rate of HSCT in first CR was observed for $MRD2_{neg}$ (32/121, 26.4%) versus $MRD2_{pos}$ groups (28/73, 38.3%, P=0.108). We highlighted a similar pattern within $MRD2_{neg}$ patients for SDAC (10/50, 20.0%) and HDAC-treated cases (22/71, 31.0%, P=0.212). Then, to estimate the benefit from HSCT in $MRD2_{neg}$ cases according to the combined model (age + treatment) category, we used the Mantel-Byar method and Simon-Makuch plots to correct for pre-

HSCT events considering HSCT as a time-dependent variable. In this analysis of DFS, the protection from relapse was not significant in COMB_1 group (P=0.64, HR=0.60, Figure S15), whereas it benefitted the COMB_2-3 patients (P=0.00433, HR=0.26, Figure S16).

Analysis of literature

Our data indicated an interaction between the predictive value of $MRD2_{neg}$ status and the intensity of treatment. We thus carried out an analysis of available literature to validate our findings. We selected a set of papers as detailed in Methods section; studies selected for analysis of DFS in $MRD2_{neg}$ cases based on treatment intensity were processed as in conventional meta-analyses and extracted MRD data are summarized in a Forest plot (Supplemental data S5). Then we performed DFS analysis after merging extracted data annotated for MRD status after two chemotherapy cycles, AML subset and treatment schedule. We selected cases that resulted negative after two chemotherapy cycles (i.e., $MRD2_{neg}$), stratifying patients according to the intensity of treatment (SDAC- vs HDAC-based regimens).

First, we carried out an analysis unselected for AML subset, and we observed a difference in DFS in favor of SDAC-treated group (P=0.014, Figure 4A). To adjust for the fact that CBF-related cases were restricted to HDAC-receiving group, we performed further analyses after excluding CBF-related cases (Figure 4B) and focusing on intermediate-risk karyotype (Figure 4C). In both analyses, SDAC-treated category displayed longer DFS (P<0.0001 and P=0.0018, respectively), a figure consistent with the findings from our dataset.

**Discussion**

MRD assessment is an essential tool for the management of patients with AML. In fact, it integrates the baseline and pre-treatment prognostic variables by conveying information on the sensitivity of the disease to treatment in the early therapeutic phases[2–5,9,12,25]. In terms of decision-making, the persistence of MRD after intensive treatment correlates with a very high-risk of disease recurrence, and patients are immediately allocated to allogeneic transplant[26]. Conversely, an undetectable MRD is far trickier to interpret and translate into a clinical decision. Although with some variability based on the methodology used, the rate of patients who relapse despite the achievement of an MRD negative status is not negligible, ranging from 20 to 40% in the first two years from CR[6,27]. The reasons for the failure of $MRD_{neg}$ to predict for maintenance of CR in such a relevant fraction of patients are not fully understood. Relapse is supposed to derive from a minor population of leukemic cells that escape MRD detection, likely due to a combination of factors. These include technical issues in the MRD measurement, as supported by the improved prognostic performance

obtained with the use of more sensitive molecular methods[28–30] or by increasing the number of parameters and cells analyzed by MFC[8,10]. Indeed, it has been shown that introduction of more stringent criteria for the limit of detection and quantitation resulted in a more accurate ability to predict patient outcome[31]. Furthermore, residual AML cells might escape detection as the result of changes in their phenotypic profile or confinement in more immature cell compartments. To address these specific issue, some investigators have proposed advancements to standard approaches, aimed at extending MRD evaluation to deviations from physiological hemopoiesis (i.e., so-called "different from normal")[7] or by highlighting abnormalities in leukemic stem cells[32,33], both strategies remaining to be validated yet, however.

Whatever the reasons, failure of $MRD_{neg}$ status to appropriately predict for maintenance of CR has relevant clinical consequences. Particularly when a relapse occurs in the first year after completion of treatment, it can be challenging to obtain a second response, outside the CBF-related subset[34]. In these patients, a negative MRD status can thus misguide the treatment planning, including the timely decision to exploit allogeneic HSCT when the disease burden is low.

Pending further methodological improvements (i.e., NGS), the clinical application of current $MRD_{neg}$ data should take into an account those limitations, and efforts should be devoted to shape the settings in which $MRD_{neg}$ predictive value is robust enough or, at the opposite, relatively weak to influence important therapeutic decisions. With such an aim in mind, we searched for interactions between MRD, baseline characteristics (age, gender, WBC, ELN stratification) and the intensity of treatment in a series of AML patients from our Institution. We focused on the predictive value of MRD2 negative status after two chemotherapy cycles (MRD2) in view of its greater capability to discriminate DFS in comparison with earlier assessment (i.e., after induction, MRD1), consistent with previous experiences[4,35].

Among baseline features, we found that younger $MRD2_{neg}$ patients (i.e., aged less than 55 y) had a significantly longer DFS (median not reached) compared to their elderly counterpart (median 25.0 months, P=0.013, HR=2.08, 95% CI 1.15-3.77). Treatment intensity, specifically the delivery of high dose ARA-C either in induction or first consolidation, apparently had a pejorative effect on the outcome of $MRD2_{neg}$ patients. This finding can be likely explained by the progressive selection of patients with unfavorable characteristics exerted by HDAC, as supported by the enrichment in ELN 2017 adverse risk category for HDAC-treated in comparison to SDAC-treated patients from post induction (19.0% vs 11.2, respectively, P=0.46) to post consolidation (16.9% vs 2.0%, P=0.014). In other words, one might surmise that the early delivery of HDAC could promote CR attainment, eventually making eligible for MRD assessment a category of patients featured by inferior chemosensitivity, that would otherwise have resulted refractory (or $MRD_{pos}$) when treated with

SDAC. HDAC could then transiently mask intrinsically chemorefractory disease, thereby correlating with a higher relapse rate than SDAC, despite both treatments resulting in MRD2 negative status.

Our findings on the effect of treatment on the significance of MRD are consistent with a previous report by Maurillo *et al*: aiming to use MRD as a biomarker for optimal ARA-C dosing, they also described lower reliability of MRD in discriminating prognosis in patients treated with HDAC compared to SDAC[36].

To validate our observations, we selected and interrogated the available literature on the interaction between treatment effect and $MRD2_{neg}$ status. The results of this analysis confirmed the pattern from our patient series, with longer DFS for patients receiving SDAC in induction and first consolidation (Figure 4).

The combination of age and treatment effects allowed to identify categories of patients with remarkably different rates of relapse, in spite of the achievement of MRD2 negativity. The group of elderly, HDAC-receiving (COMB_2-3) patients displayed a DFS rate of around 50% at 3 years. This figure challenges the usual implications of $MRD_{neg}$ status, as current ELN guidelines recommend consolidation with allogeneic HSCT as the preferred post CR option for patients with an estimated relapse risk exceeding 35-40%[1,37]. Of note, this category of patients showed the greatest benefits from the delivery of HSCT in CR1, as emerging from a specific analysis by Mantel-Byar method (Figure S16). This finding is consistent with the established influence of disease burden, usually estimated by MRD, on the efficacy of the antileukemic effect of HSCT[26,38,39]. Indeed, it is plausible that, although both characterized by $MRD2_{neg}$ status, COMB_1 achieved a more profound level of response than COMB_2-3 patients, likely explaining the difference in outcome in the post-HSCT setting.

The longer DFS for SDAC- vs HDAC-treated $MRD2_{neg}$ patients kept in intermediate-risk tier, as assessed by karyotype and ELN (Figure S13), the category where MRD2 assessment is generally employed for the clinical management of AML patients.

These data emphasize some of the current limitations of MRD assessment. Although MRD represents the strongest predictive parameter for long-term survival overall, in some settings it actually lacks the robustness needed to support key clinical decisions. In particular, the translation of an $MRD2_{neg}$ status into a consolidation program including or omitting HSCT, should take into account other contributing variables. At this regard, herein we provide evidence that age and treatment intensity may help to effectively delineate settings (younger age, SDAC-treated), where the application of $MRD2_{neg}$ data to support clinical decisions is justified based on its correlation with an excellent outcome, unlike in others instances (older aged patients who were HDAC-treated)

where it did not add to the baseline stratification and should not be used as key decision-making parameter. As with the dose of ARA-C, other treatment variables (i.e., GO, FLT3 inhibitors) may also have an effect on MRD, an issue that deserves to be evaluated in large series of patients.

Our work has some limitations to be acknowledged. We recognize the limits of a retrospective study with a wide enrollment period time, with changes in risk assessment and treatment allocation (in particular with regards to HSCT), although we checked for an impact for year of diagnosis on OS (Figure S1). Another point concerns the limited number of patients in the adverse risk category according to karyotype and ELN, preventing specific analyses. This finding is clearly due to the selection of patients who had a response to chemotherapy, at least initially, and in any case represents the disease subset where MRD is generally not crucial in clinical decisions. The lack of validation of our results in an independent cohort is a further weakness, that we tried to address with the analysis of data extracted from the available literature which confirmed our findings on the impact of treatment.

In conclusion, we showed a variable reliability of MRD in different AML settings, as defined by age and intensity of treatment. In addition to underscoring the need for further improvement of MRD approaches, these findings call for a reasoned application of MRD information, as currently available, to modulate consolidation therapy on adequately estimated relapse rates.

**References**


1. Döhner H, Wei AH, Appelbaum FR, et al. Diagnosis and Management of AML in Adults: 2022 ELN Recommendations from an International Expert Panel. Blood. 2022;140(12):1345-1377.

2. Miguel JFS, Martínez A, Macedo A, et al. Immunophenotyping investigation of minimal residual disease is a useful approach for predicting relapse in acute myeloid leukemia patients. Blood. 1997;90(6):2465-2470.

3. Venditti A, Buccisano F, Poeta DG, et al. Level of minimal residual disease after consolidation therapy predicts outcome in acute myeloid leukemia. Blood. 2000;96(12):3948-3952.

4. Terwijn M, Putten WL van, Kelder A, et al. High prognostic impact of flow cytometric minimal residual disease detection in acute myeloid leukemia: data from the HOVON/SAKK AML 42A study. J Clin Oncol. 2013;31(31):3889-3897.

5. 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.

6. Short NJ, Zhou S, Fu C, et al. Association of Measurable Residual Disease With Survival Outcomes in Patients With Acute Myeloid Leukemia. JAMA Oncol. 2020;6(12):1890-1899.

7. Heuser M, Freeman SD, Ossenkoppele GJ, et al. 2021 Update Measurable Residual Disease in Acute Myeloid Leukemia: European LeukemiaNet Working Party Consensus Document. Blood. 2021;138(26):2753-2767.

8. Buccisano F, Maurillo L, Principe MID, 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.

9. Freeman SD, Hills RK, Virgo P, et al. Measurable Residual Disease at Induction Redefines Partial Response in Acute Myeloid Leukemia and Stratifies Outcomes in Patients at Standard Risk Without NPM1 Mutations. J Clin Oncol. 2018;36(15):1486-1497.

10. Tettero JM, Freeman S, Buecklein V, et al. Technical Aspects of Flow Cytometry-based Measurable Residual Disease Quantification in Acute Myeloid Leukemia: Experience of the European LeukemiaNet MRD Working Party. Hemasphere. 2021;6(1):e676.

11. Zhu -H H, Zhang -H X, Qin -Z Y, et al. MRD-directed risk stratification treatment may improve outcomes of t(8;21) AML in the first complete remission: results from the AML05 multicenter trial. Blood. 2013;121(20):4056-4062.

12. 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.

13. Buccisano F, Maurillo L, Piciocchi A, et al. Minimal residual disease negativity in elderly patients with acute myeloid leukemia may indicate different postremission strategies than in younger patients. Ann Hematol. 2015;94(8):1319-1326.

14. Jentzsch M, Grimm J, Bill M, et al. Clinical Value of the Measurable Residual Disease Status within the ELN2017 Risk Groups in AML Patients undergoing Allogeneic Stem Cell Transplantation. Am J Hematol. 2021;96(7):E237-E239.

15. Burnett AK, Goldstone A, Hills RK, et al. Curability of Patients With Acute Myeloid Leukemia Who Did Not Undergo Transplantation in First Remission. J Clin Oncol. 2013;31(10):1293-1301.

16. Castaigne S, Chevret S, Archimbaud E, et al. Randomized comparison of double induction and timed-sequential induction to a "3 + 7" induction in adults with AML: long-term analysis of the Acute Leukemia French Association (ALFA) 9000 study. Blood. 2004;104(8):2467-2474.

17. Bassan R, Intermesoli T, Masciulli A, et al. Randomized trial comparing standard vs sequential high-dose chemotherapy for inducing early CR in adult AML. Blood Adv. 2019;3(7):1103-1117.

18. Mannelli F, Bencini S, Piccini M, et al. Multilineage Dysplasia as Assessed by Immunophenotype in Acute Myeloid Leukemia: A Prognostic Tool in a Genetically Undefined Category. Cancers. 2020;12(11):3196.

19. Mannelli F, Gianfaldoni G, Bencini S, et al. Early peripheral blast cell clearance predicts minimal residual disease status and refines disease prognosis in acute myeloid leukemia. Am J Hematol. 2020;95(11):1304-1313.

20. Schuurhuis GJ, Heuser M, Freeman S, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. Blood. 2018;131(12):1275-1291.

21. Gabert J, Beillard E, Velden VHJ van der, 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.

22. Grimwade D, Hills RK, Moorman AV, et al. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. Blood. 2010;116(3):354-365.

23. 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.

24. Guyot P, Ades A, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. Bmc Med Res Methodol. 2012;12(1):9.

25. Balsat M, Renneville A, Thomas X, et al. Postinduction Minimal Residual Disease Predicts Outcome and Benefit From Allogeneic Stem Cell Transplantation in Acute Myeloid Leukemia With NPM1 Mutation: A Study by the Acute Leukemia French Association Group. J Clin Oncol. 2016;35(2):185-193.

26. Buccisano F, Maurillo L, Piciocchi A, et al. Pre-transplant persistence of minimal residual disease does not contraindicate allogeneic stem cell transplantation for adult patients with acute myeloid leukemia. Bone Marrow Transpl. 2017;52(3):473-475.

27. Paiva B, Vidriales M-B, Sempere A, et al. Impact of measurable residual disease by decentralized flow cytometry: a PETHEMA real-world study in 1076 patients with acute myeloid leukemia. Leukemia. 2021;35(8):2358-2370.

28. Pastore F, Levine RL. Next-Generation Sequencing and Detection of Minimal Residual Disease in Acute Myeloid Leukemia: Ready for Clinical Practice? JAMA. 2015;314(8):778-780.

29. Klco JM, Miller CA, Griffith M, et al. Association Between Mutation Clearance After Induction Therapy and Outcomes in Acute Myeloid Leukemia. JAMA. 2015;314(8):811-822.

30. Jongen-Lavrencic M, Grob T, Hanekamp D, et al. Molecular Minimal Residual Disease in Acute Myeloid Leukemia. N Engl J Med 2018;378(13):1189-1199.

31. Buccisano F, Palmieri R, Piciocchi A, et al. Clinical relevance of an objective flow cytometry approach based on limit of detection and limit of quantification for measurable residual disease assessment in acute myeloid leukemia. A post-hoc analysis of the GIMEMA AML1310 trial. Haematologica. 2022;107(12):2823-2833.

32. Li S-Q, Xu L-P, Wang Y, et al. An LSC-based MRD assay to complement the traditional MFC method for prediction of AML relapse: a prospective study. Blood. 2022;140(5):516-520.

33. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJM, et al. A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. Leukemia. 2016;30(2):439-446.

34. Breems DA, Putten WLJV, Huijgens PC, et al. Prognostic Index for Adult Patients With Acute Myeloid Leukemia in First Relapse. J Clin Oncol. 2005;23(9):1969-1978.

35. Maurillo L, Buccisano F, Principe M, et al. Toward Optimization of Postremission Therapy for Residual Disease–Positive Patients With Acute Myeloid Leukemia. J Clin Oncol. 2008;26(30):4944-4951.

36. 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.

37. Cornelissen JJ, Blaise D. Hematopoietic stem cell transplantation for patients with AML in first complete remission. Blood. 2016;127(1):62-70.

38. Walter RB, Gooley TA, Wood BL, et al. Impact of Pretransplantation Minimal Residual Disease, As Detected by Multiparametric Flow Cytometry, on Outcome of Myeloablative Hematopoietic Cell Transplantation for Acute Myeloid Leukemia. J Clin Oncol. 2011;29(9):1190-1197.

39. Hourigan CS, Dillon LW, Gui G, et al. Impact of Conditioning Intensity of Allogeneic Transplantation for Acute Myeloid Leukemia With Genomic Evidence of Residual Disease. J Clin Oncol. 2019;38(12):1273-1283.

**Table 1.** Characteristics of patients according to treatment group after induction and first consolidation cycle.

| | **Overall n = 194** | **SDAC n = 70 (36.1%)** | **HDAC n = 124 (63.9%)** | ***P* value** | |
|---|---|---|---|---|---|
| **Age,** median (range) | 55 (16-73) | 55 (22–70) | 53 (16–73) | 0.75 | |
| **WBC,** x$10^9$/L, median (range) | 14.8 (0.6-435.0) | 14.9 (0.6–191.0) | 13.6 (0.9–435.0) | 0.90 | |
| **Hb,** g/dL, median (range) | 9.2 (3.9-14.9) | 9.2 (5.4-14.5) | 9.0 (3.9-14.9) | 0.25 | |
| **Plt,** x$10^9$/L, median (range) | 53 (1-281) | 47 (10-152) | 57 (1-281) | 0.053 | |
| **Peripheral blasts,** %, median (range) | 53 (0-100) | 58 (0-100) | 52 (0-98) | 0.30 | |
| **Secondary AML,** n (%) | 12 (6.2%) | 3 (4.3) | 9 (7.3) | 0.54 | |
| **Karyotype,** n (%) | | | | | |
| *Favorable* | 33 (17.0) | 13 (18.6) | 20 (16.1) | 0.69 | |
| *Normal* | 110 (56.7) | 41 (58.6) | 69 (55.6) | 0.76 | |
| *Intermediate, non-normal* | 26 (13.4) | 9 (12.9) | 17 (13.7) | 1.0 | 0.38 |
| *Adverse* | 15 (7.7) | 2 (2.9) | 13 (10.5) | 0.09 | |
| *Lack of growth* | 10 (5.2) | 5 (7.1) | 5 (4.0) | 0.50 | |
| **Molecular genetics,** n (%) | | | | | |
| *NPM1-mutated* | 80 (41.2) | 39 (55.7) | 47 (37.9) | 0.23 | |
| *FLT3-ITD* | 41 (21.1) | 16 (22.9) | 25 (20.2) | 0.71 | |
| *CEBPA-bZIP* | 10 (5.2) | 4 (5.7) | 6 (4.8) | 0.74 | |
| **ELN 2010 risk groups,** n (%) | | | | | |
| *Favorable* | 90 (46.4) | 39 (55.7) | 51 (41.1) | 0.06 | |
| *Intermediate-1* | 80 (41.2) | 27 (38.6) | 53 (42.7) | 0.64 | **0.01** |
| *Intermediate-2* | 8 (4.1) | 2 (2.9) | 6 (4.8) | 0.71 | |
| *Adverse* | 16 (8.2) | 2 (2.9) | 14 (11.3) | 0.06 | |
| **ELN 2017 risk group,** n (%) | | | | | |
| *Favorable* | 102 (52.5) | 41 (58.7) | 61 (49.2) | 0.23 | |
| *Intermediate* | 62 (32.0) | 19 (27.1) | 43 (34.7) | 0.33 | 0.05 |
| *Adverse* | 24 (12.4) | 5 (7.1) | 19 (15.3) | 0.11 | |
| *Not assessable* | 6 (3.1) | 5 (7.1) | 1 (0.8) | | |

Differences between treatment groups were evaluated using Mann-Whitney test for continuous variables and Fisher exact tests or $\chi 2$ for categorical variables. Values in bold are statistically significant ($P < 0.05$). Abbreviations: WBC, white blood cells; Hb, hemoglobin, Plt, platelets; ELN, European Leukemia Net; SDAC: standard dose cytarabine (both in induction and consolidation cycle); HDAC: high dose cytarabine (at least in one cycle, induction or first consolidation).

**Table 2.** Characteristics of patients $MRD2_{neg}$ according to treatment group after induction and first consolidation cycle.

| | **Overall $MRD2_{neg}$ n = 121** | **$MRD2_{neg}$ SDAC n = 50 (41.3%)** | **$MRD2_{neg}$ HDAC n = 71 (58.7%)** | ***P* value** | |
|---|---|---|---|---|---|
| **Age,** median (range) | 55 (22-73) | 55 (22-69) | 55 (22-73) | 0.23 | |
| **WBC,** $x10^9$/L, median (range) | 8.9 (0.6-380.0) | 11.8 (0.6-191.0) | 7.5 (0.9-380.0) | 0.30 | |
| **Hb,** g/dL, median (range) | 9.3 (3.9-14.1) | 9.3 (5.4-13.3) | 9.3 (3.9-14.1) | 0.55 | |
| **Plt,** $x10^9$/L, median (range) | 56 (9-271) | 47 (10-152) | 71 (9-271) | **0.03** | |
| **Peripheral blasts,** %, median (range) | 47 (0-100) | 57 (0-100) | 40 (0-98) | 0.06 | |
| **Secondary AML,** n (%) | 10 (8.3) | 7 (14.0) | 3 (4.2) | 0.09 | |
| **Karyotype,** n (%) | | | | | |
| *Favorable* | 13 (10.7) | 8 (16.0) | 5 (7.0) | 0.14 | |
| *Normal* | 72 (59.5) | 31 (62.0) | 41 (57.7) | 0.71 | |
| *Intermediate, non-normal* | 19 (15.7) | 6 (12.0) | 13 (18.3) | 0.45 | 0.148 |
| *Adverse* | 9 (7.5) | 1 (2.0) | 8 (11.3) | 0.07 | |
| *Lack of growth* | 8 (6.6) | 4 (8.0) | 4 (5.6) | 0.72 | |
| **Molecular genetics,** n (%) | | | | | |
| *NPM1-mutated* | 50 (41.3) | 25 (50.0) | 25 (35.2) | 0.13 | |
| *FLT3-ITD* | 20 (16.5) | 9 (18.0) | 11 (15.5) | 0.19 | |
| *CEBPA-bZIP* | 10 (8.3) | 4 (8.0) | 6 (8.5) | 1.0 | |
| **ELN 2010 risk groups,** n (%) | | | | | |
| *Favorable* | 55 (45.5) | 30 (60.0) | 25 (35.2) | **0.009** | |
| *Intermediate-1* | 50 (41.3) | 17 (34.0) | 33 (46.5) | 0.19 | **0.003** |
| *Intermediate-2* | 6 (5.0) | 2 (4.0) | 4 (5.6) | 1.0 | |
| *Adverse* | 10 (8.2) | 1 (2.0) | 9 (12.7) | **0.045** | |
| **ELN 2017 risk group,** n (%) | | | | | |
| *Favorable* | 62 (51.3) | 32 (64.0) | 30 (42.2) | **0.026** | |
| *Intermediate* | 42 (34.7) | 13 (24.0) | 29 (40.8) | 0.12 | **0.001** |
| *Adverse* | 13 (10.7) | 1 (2.0) | 12 (16.9) | **0.014** | |
| *Not assessable* | 4 (3.3) | 4 (10.0) | 0 (-) | **0.027** | |

Differences between treatment groups were evaluated using Mann-Whitney test for continuous variables and Fisher exact tests or $\chi^2$ for categorical variables. Values in bold are statistically significant ($P < 0.05$). Abbreviations: WBC, white blood cells; Hb, hemoglobin, Plt, platelets; ELN, European Leukemia Net; SDAC: standard dose cytarabine (both in induction and consolidation cycle); HDAC: high dose cytarabine (at least in one cycle, induction or first consolidation).

**Figure legends.**

**Figure 1. Kinetics of MRD according to time-point of assessment.** Flow of data for MRD status (negative, $MRD_{neg}$ / positive, $MRD_{pos}$) at post induction (MRD1) and post consolidation (MRD2) time-points. The chemotherapy regimens and the number of patients treated with each scheme are enlisted for induction (pink panel) and consolidation (blue panel) cycle. The cycles were classified according to the dose of cytarabine as standard dose (SDAC) or high dose (HDAC). Treatment details for each chemotherapy scheme are provided in Supplemental file. Abbreviations: GO: gemtuzumab ozogamicin; Mido: midostaurin; CPX: CPX-351; ICE: idarubicin, cytarabine, etoposide; HDS: high-dose sequential; Ida: idarubicin; HDAC 10: high dose cytarabine 10 $gr/m^2$; HDAC 1,3,5: high dose cytarabine days 1, 3, and 5; DIA: intermediate dose cytarabine.

**Figure 2. Analysis of survival according to MRD2 status in different treatment groups.** Disease-free survival (A-C) and overall survival (B-D) according to MRD status after first consolidation (MRD2) in patients treated with standard dose cytarabine in both induction and consolidation (A-B) or with at least one cycle including high dose cytarabine (C-D). The curves of patients with $MRD2_{neg}$ status are depicted in blue; the curves of patients with $MRD2_{pos}$ status are depicted in red. Median survival estimates are reported in months. NR, not reached; HR, hazard ratio; 95% CI, confidence interval; NA, not available.

**Figure 3. Analysis of major outcomes for $MRD2_{neg}$ patients according to the combined model.** Disease-free survival (A-C) and overall survival (B-D) according to the combined model including age range and treatment intensity in $MRD2_{neg}$ patients: the curves of patients aged less than 55y and treated with standard dose cytarabine in both induction and consolidation (COMB_1) are depicted in blue. In the upper panels, the curves of patients aged more or equal to 55y and treated with high dose cytarabine in at least one of the first two cycles (COMB_3) are depicted in red; the remainder patients (COMB_2) are depicted in grey. In the lower panels, COMB_2 and COMB_3 categories are grouped, and their survival curves depicted in red. Median survival estimates are reported in months. NR, not reached; HR, hazard ratio; 95% CI, confidence interval; NA, not available.

**Figure 4. Analysis of disease-free survival according to treatment intensity in data extracted from literature.** Disease-free survival in $MRD2_{neg}$ cases according to treatment intensity, SDAC (blue curve) versus HDAC (red curve) in different AML subsets: (A) unselected; (B) after exclusion of CBF-related AML; (C) intermediate-risk karyotype. Below each Kaplan-Meier curve,

the papers from which data are extracted for each analysis are listed by first author name and year of publication.

Figure 1

**INDUCTION**

- 3+7 (± GO-Mido, CPX) # 43
- ICE # 100
- HDS # 29
- Ida-HDAC # 44
- HDAC 10 # 8

**CONSOLIDATION**

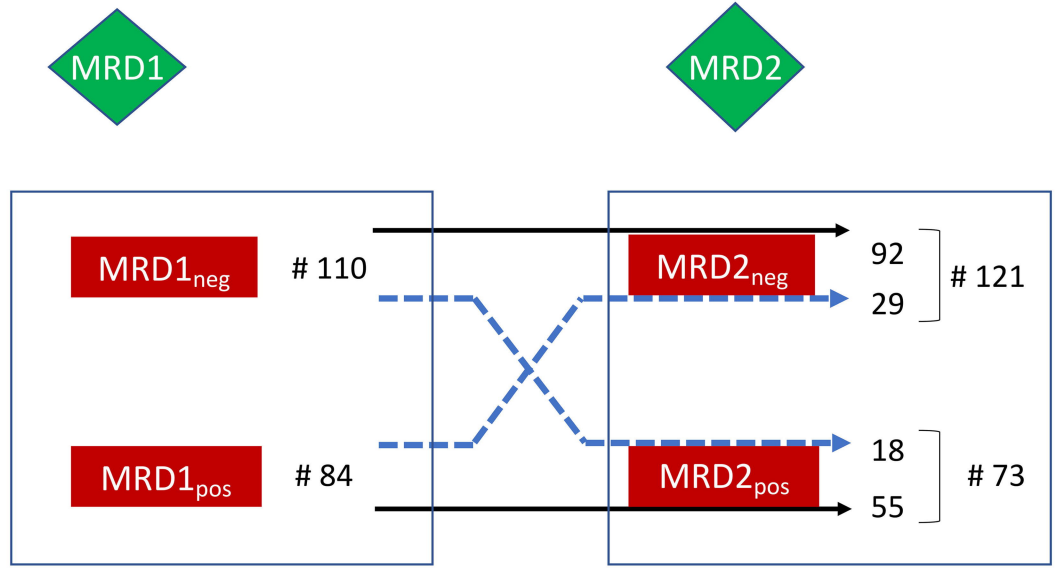
- HDAC 1,3,5 # 77
- A8-A10 # 37
- IC # 101
- Ida-HDAC # 4
- DIA-GO # 5

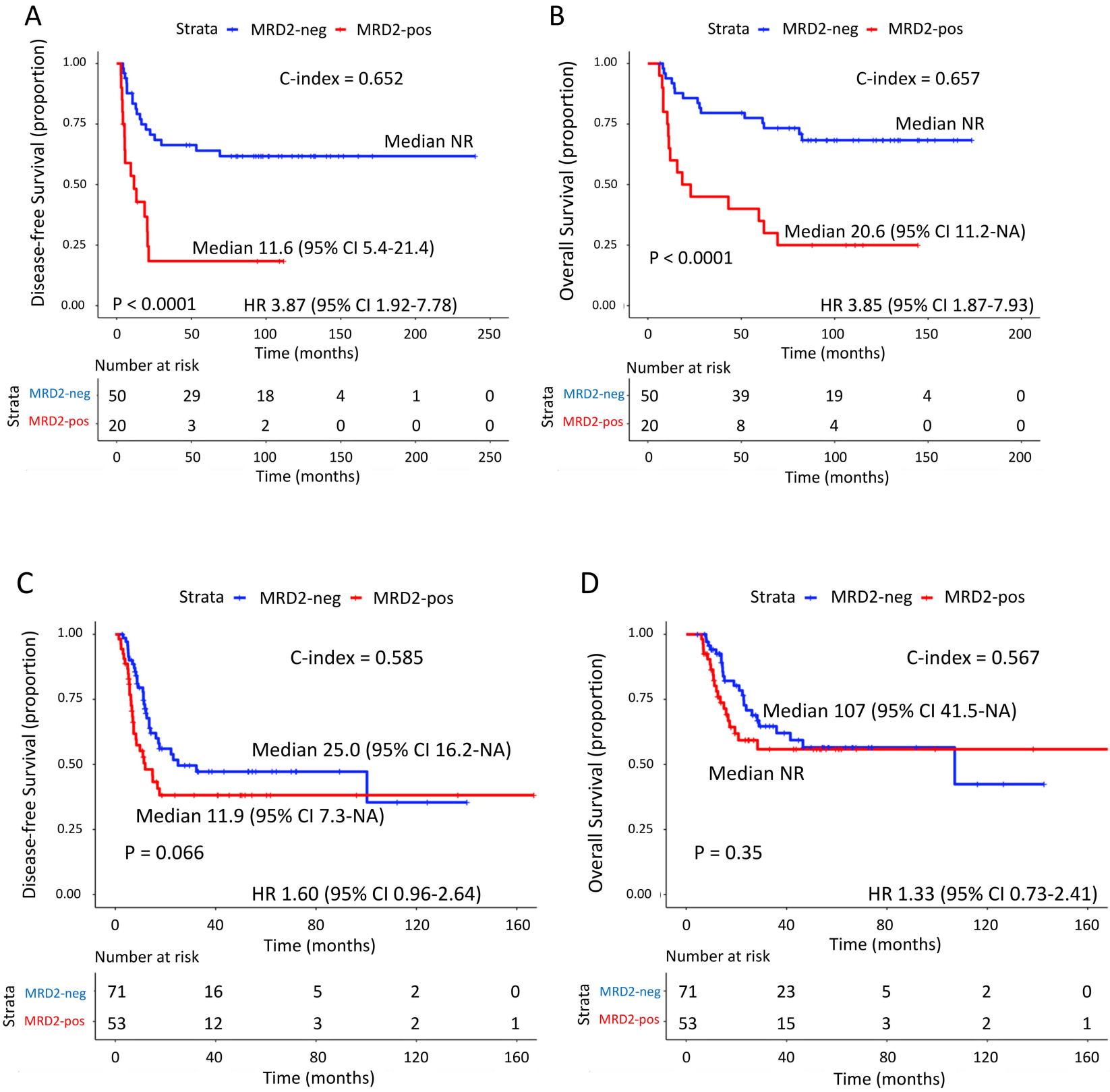
Treatment combinations:

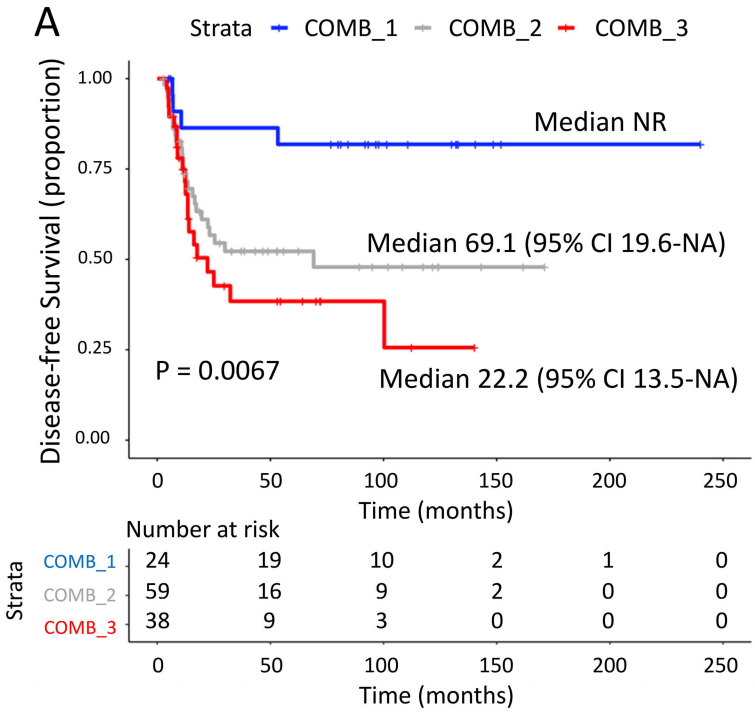
- ✓ SDAC/SDAC # 70 (36.1%)
- ✓ HDAC-containing # 124 (63.9%)
  - SDAC-HDAC # 77
  - HDAC-HDAC # 47

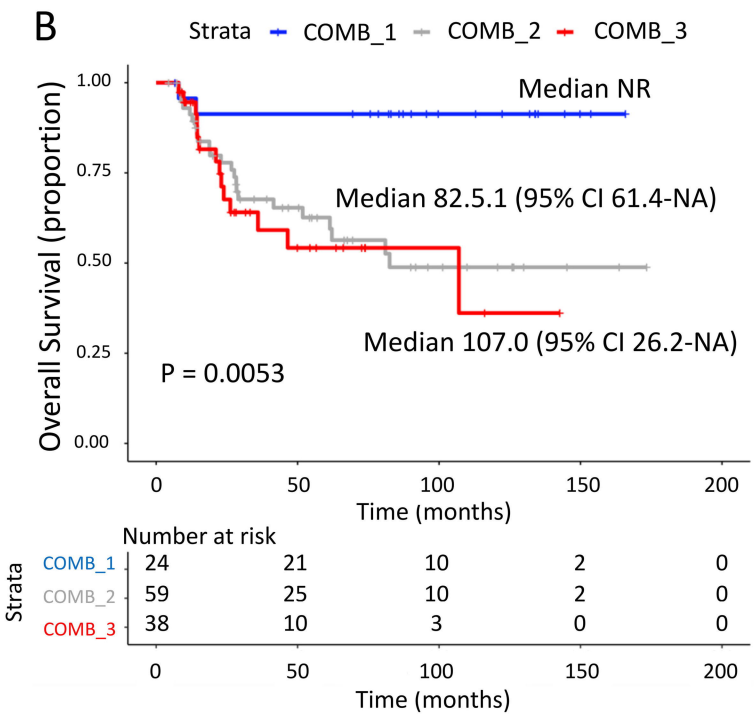
MRD methods:

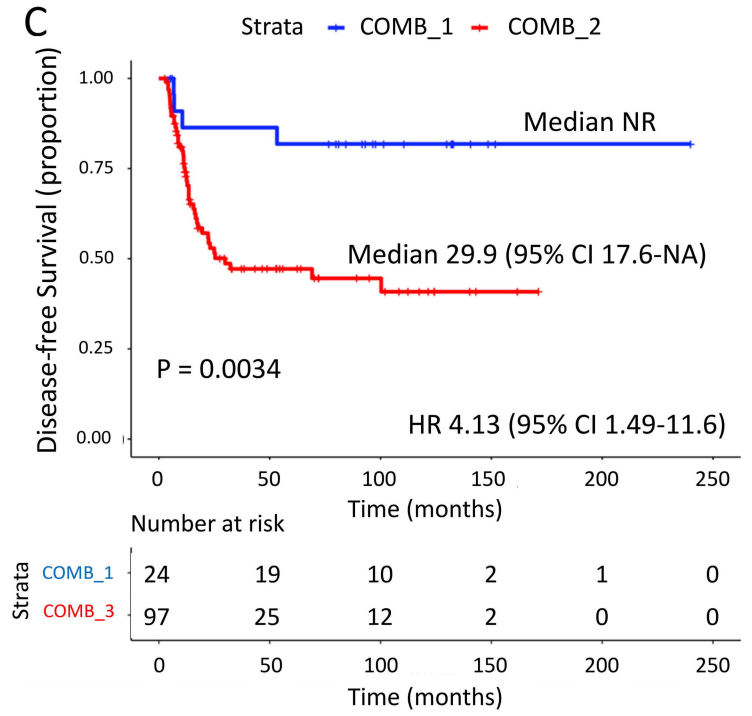
- ✓ CBF/NPM1+: PCR-based
- ✓ Other: MFC LAIP-based

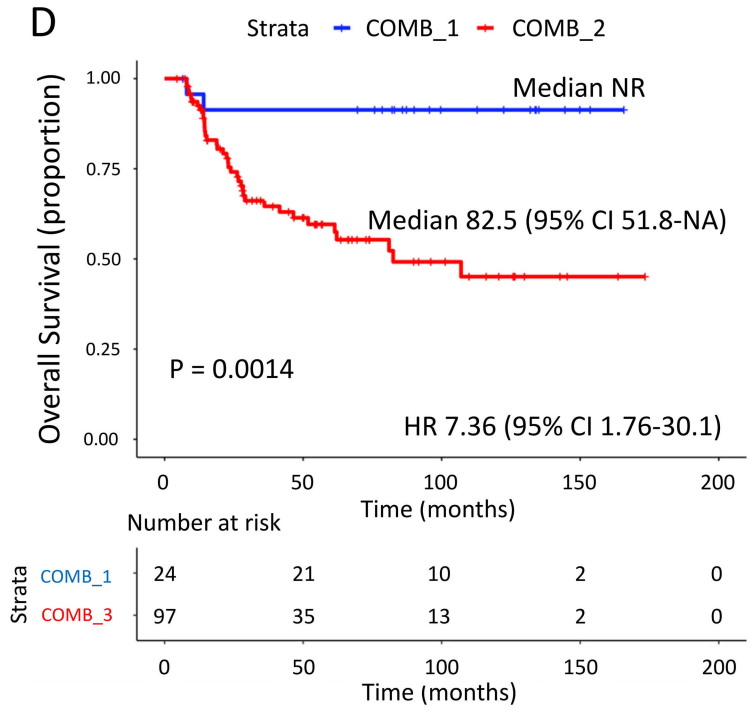



A
Strata MRD2-neg MRD2-pos
Disease-free Survival (proportion)
C-index = 0.652
Median NR
Median 11.6 (95% CI 5.4-21.4)
P < 0.0001
HR 3.87 (95% CI 1.92-7.78)
Time (months)
Number at risk
MRD2-neg 50 29 18 4 1 0
MRD2-pos 20 3 2 0 0 0
B
Strata MRD2-neg MRD2-pos
Overall Survival (proportion)
C-index = 0.657
Median NR
Median 20.6 (95% CI 11.2-NA)
P < 0.0001
HR 3.85 (95% CI 1.87-7.93)
Time (months)
Number at risk
MRD2-neg 50 39 19 4 0
MRD2-pos 20 8 4 0 0
C
Strata MRD2-neg MRD2-pos
Disease-free Survival (proportion)
C-index = 0.585
Median 25.0 (95% CI 16.2-NA)
Median 11.9 (95% CI 7.3-NA)
P = 0.066
HR 1.60 (95% CI 0.96-2.64)
Time (months)
Number at risk
MRD2-neg 71 16 5 2 0
MRD2-pos 53 12 3 2 1
D
Strata MRD2-neg MRD2-pos
Overall Survival (proportion)
C-index = 0.567
Median 107 (95% CI 41.5-NA)
Median NR
P = 0.35
HR 1.33 (95% CI 0.73-2.41)
Time (months)
Number at risk
MRD2-neg 71 23 5 2 0
MRD2-pos 53 15 3 2 1

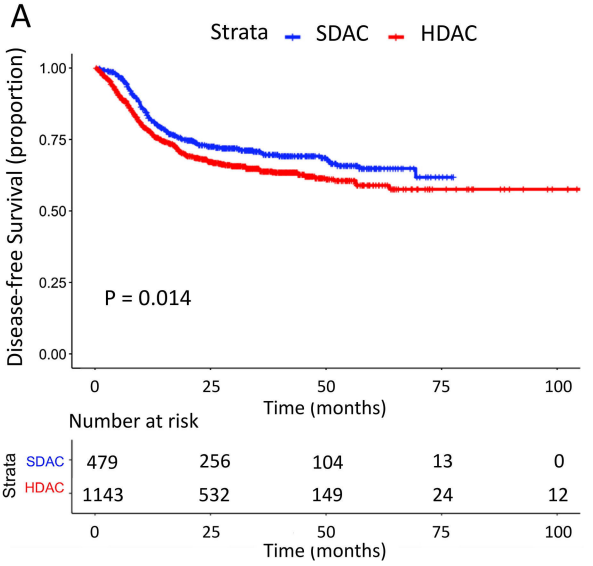

A
Strata COMB_1 COMB_2 COMB_3
Disease-free Survival (proportion)
1.00
0.75
0.50
0.25
0.00
Median NR
Median 69.1 (95% CI 19.6-NA)
P = 0.0067
Median 22.2 (95% CI 13.5-NA)
0 50 100 150 200 250
Time (months)
Number at risk
Strata
COMB_1 24 19 10 2 1 0
COMB_2 59 16 9 2 0 0
COMB_3 38 9 3 0 0 0
0 50 100 150 200 250
Time (months)


B
Strata COMB_1 COMB_2 COMB_3
Overall Survival (proportion)
1.00
0.75
0.50
0.25
0.00
Median NR
Median 82.5.1 (95% CI 61.4-NA)
Median 107.0 (95% CI 26.2-NA)
P = 0.0053
0 50 100 150 200
Time (months)
Number at risk
Strata
COMB_1 24 21 10 2 0
COMB_2 59 25 10 2 0
COMB_3 38 10 3 0 0
0 50 100 150 200
Time (months)


C
Strata COMB_1 COMB_2
Disease-free Survival (proportion)
1.00
0.75
0.50
0.25
0.00
Median NR
Median 29.9 (95% CI 17.6-NA)
P = 0.0034
HR 4.13 (95% CI 1.49-11.6)
0 50 100 150 200 250
Time (months)
Number at risk
Strata
COMB_1 24 19 10 2 1 0
COMB_3 97 25 12 2 0 0
0 50 100 150 200 250
Time (months)

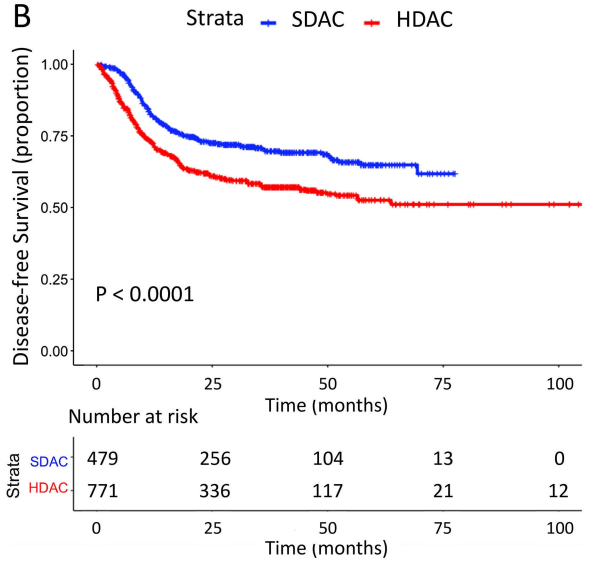

D
Strata COMB_1 COMB_2
Overall Survival (proportion)
1.00
0.75
0.50
0.25
0.00
Median NR
Median 82.5 (95% CI 51.8-NA)
P = 0.0014
HR 7.36 (95% CI 1.76-30.1)
0 50 100 150 200
Time (months)
Number at risk
Strata
COMB_1 24 21 10 2 0
COMB_3 97 35 13 2 0
0 50 100 150 200
Time (months)

A



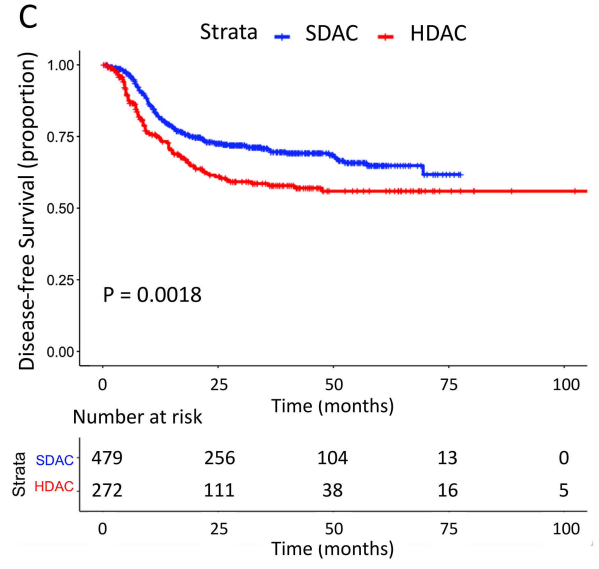

Papers:
- ✓ Buccisano 2010
- ✓ Chou 2010
- ✓ Jourdan 2013
- ✓ Narimatsu 2008
- ✓ Onecha 2019
- ✓ Ravandi 2017
- ✓ Shayegi 2013
- ✓ Terwjin 2012
- ✓ Terwjin 2013
- ✓ Zhang 2013
- ✓ Bataller 2021
- ✓ Wei 2021
- ✓ Freeman 2018
- ✓ Hubmann 2014
- ✓ Ivey 2016
- ✓ Jongen-Lavrencic 2020
- ✓ Kapp-Schwoerer 2016
- ✓ Willekens 2016
- ✓ Zeijlemaker 2015

B




Papers:
- ✓ Buccisano 2010
- ✓ Chou 2010
- ✓ Onecha 2019
- ✓ Ravandi 2017
- ✓ Shayegi 2013
- ✓ Terwjin 2012
- ✓ Terwjin 2013
- ✓ Freeman 2018
- ✓ Hubmann 2014
- ✓ Ivey 2016
- ✓ Jongen-Lavrencic 2018
- ✓ Kapp-Schwoerer 2020
- ✓ Zeijlemaker 2015

C




Papers:
- ✓ Chou 2010
- ✓ Onecha 2019
- ✓ Shayegi 2013
- ✓ Terwjin 2013
- ✓ Bataller 2021
- ✓ Freeman 2018
- ✓ Hubmann 2014
- ✓ Ivey 2016
- ✓ Kapp-Schwoerer 2016

## Supplemental data

# The effect of age and treatment on the predictive value of measurable residual disease: implications for the clinical management of adult patients with acute myeloid leukemia

Francesco Mannelli[1,2], Matteo Piccini[1], Sara Bencini[3], Giacomo Gianfaldoni[1], Benedetta Peruzzi[3], Roberto Caporale[3], Barbara Scappini[1], Laura Fasano[1], Elisa Quinti[1], Gaia Ciolli[1], Andrea Pasquini[1], Francesca Crupi[1], Sofia Pilerci[1], Fabiana Pancani[1,2], Leonardo Signori[1,2], Danilo Tarantino[1,2], Chiara Maccari[1,2], Vivian Paradiso[1], Francesco Annunziato[3], Paola Guglielmelli[1], Alessandro M. Vannucchi[1]

[1]SOD Ematologia, Università di Firenze, AOU Careggi, Firenze, Italy; [2]Centro Ricerca e Innovazione Malattie Mieloproliferative (CRIMM), AOU Careggi, Firenze, Italy; [3]Centro Diagnostico di Citofluorimetria e Immunoterapia, AOU Careggi, Firenze, Italy

# TABLE OF CONTENTS

## Supplemental Materials and Methods

### S1. Treatment protocols

Protocol-1: since April 2004 to March 2007, patients received induction according to standard-dose cytarabine (SDAC) based course, namely "3+7" (Cytarabine 100 mg/sqm bid on days 1-7; Idarubicin 12 mg/sqm on days 1-3). From 2006 on, etoposide 100 mg/sqm on days 1-5 was added (ICE course). High-dose cytarabine (HDAC) 1, 3, 5 (3000 mg/sqm bid on days 1, 3, 5) was used as first consolidation in patients aged < 61 years attaining complete remission (CR) after ICE. Patients with persistent disease (*i.e.*, > 5% BM blasts at hematopoietic recovery) after first course received a salvage regimen (Ida-HDAC). In an intention-to-treat approach, patients aged < 55 years with high-risk karyotype, *FLT3*-ITD or adverse clinical features (secondary AML, CR after second course, hyperleukocytosis) were assigned to undergo allogeneic stem cell transplantation (SCT) from matched related or unrelated donor. Patients with intermediate cytogenetic risk in the absence of *FLT3*-ITD and adverse clinical features were allocated to allogeneic SCT if a related donor was available. Autologous SCT was offered to patients aged < 61 y with low-risk cytogenetics, intermediate-risk cytogenetics without sibling donor and high-risk disease not eligible to allogeneic SCT. Peripheral blood (PB) stem cells for autologous SCT were collected after a mobilization course (Cytarabine 500 mg/sqm bid on days 1-6; Daunorubicin 50 mg/sqm on days 4-6). Patients who failed mobilization received two additional courses with high dose cytarabine.

Protocol-2: since April 2007 to April 2014, patients were treated according to Northern Italy Leukemia Group (NILG) AML 02-06 protocol. Until March 2012, patients were recruited within the NILG AML 02/06 trial [(ClinicalTrials.gov Identifier: NCT00495287; reference: Bassan R, et al; Blood Adv. 2019;3(7):1103–1117)]. From April 2012, after closure of NILG AML 02/06 trial, patients were treated according to the standard arm provided by the protocol. The protocol provided a randomization at induction between a standard ICE induction *versus* an experimental intensified one. Patients aged > 65 y were treated according to standard arm. Upon CR achievement, patients received standard doses cytarabine consolidation and were divided into standard and high-risk cases (SR, HR): SR: favorable or intermediate risk cytogenetics (according to SWOG criteria) without any adverse clinical factor (secondary AML, *FLT3*-ITD, CR after cycle 2, persistence of pre-existing cytogenetic abnormality despite morphological CR; total WBC count >50 $x10^9$/L); HR: all non-SR cases. HR patients were assigned to undergo allogeneic SCT. Provided sufficient CD34+ cells were previously collected (>$2x10^6$/kg) upon recovery from high doses cytarabine, SR patients and HR patients excluded from allo-SCT and aged 65 years or less were randomized between autologous SCT and high doses consolidation therapy (R2). HR/SR patients unable to be randomized in R2 because of inadequate blood stem cell yield received intermediate-dose consolidation. Patients randomized to experimental arm were excluded from outcome analysis.

Protocol-3: since May 2014 to April 2017, patients received induction according to Ida-FLA course, (Cytarabine 2000 mg/sqm on days 1-4; Fludarabine 30 mg/sqm on days 1-4; Idarubicin 10 mg/sqm on days 2-4). High-dose Cytarabine (3000 mg/sqm bid days 1, 3, 5) was used as first consolidation in patients aged < 61 years attaining complete remission (CR). Patients with persistent disease (*i.e.* > 5% BM blasts at hematopoietic recovery) after first course received a salvage regimen (Clofarabine-based). In post CR phase, patients were stratified according to European Leukemia Net 2010 guidelines [reference: Döhner H, et al; Blood. 2010;115(3):453–474]. Patients in adverse-risk

category were allocated to allogeneic HSCT from matched related or unrelated donor. Patients in intermediate category were allocated to allogeneic SCT if a related donor was available. Patients in favorable-risk ELN category and high-risk disease not eligible to allogeneic SCT received up to two additional courses with high dose cytarabine.

Protocol-4: since 2017, patients harboring *FLT3* mutations received induction according to "3+7" scheme (Cytarabine 200 mg/sqm intravenous continuous infusion on days 1-7; Daunorubicin 60 mg/sqm on days 1-3) + Midostaurin 50 mg bid orally on days 8-21. High-dose Cytarabine (3000 mg/sqm bid days 1, 3, 5) + Midostaurin 50 mg bid orally on days 8-21 was used as first consolidation in patients aged < 61 years attaining complete remission (CR) [reference: Stone R, New Engl J Med. 2017;377, 454]. In post CR phase, patients were stratified according to European Leukemia Net 2017 guidelines [reference: Döhner H, et al; Blood. 2022;140 (12): 1345]. Patients in adverse-risk category were allocated to allogeneic HSCT from matched related or unrelated donor. Patients in intermediate category were allocated to allogeneic SCT if a related donor was available. Patients in favorable-risk ELN category and high-risk disease not eligible to allogeneic SCT received up to two additional courses with high dose cytarabine.

Protocol-5: since 2017, elderly patients (>60 y) diagnosed with AML with myelodysplasia-related changes received induction with CPX-351 100 U/sqm intravenously on days 1, 3, 5. For patients in CR after induction, consolidation treatment provided up to two cycles of CPX-351 65 U/sqm intravenously on days 1, 3 [reference: Lancet J, et al; JCO. 2016; 36: 2684]. If eligible, patients were allocated to allogeneic HSCT from matched related or unrelated donor.

Protocol-6: since 2017, patients diagnosed with core binding factor (CBF) related AML received induction according to "3+7" scheme (Cytarabine 200 mg/sqm intravenous continuous infusion on days 1-7; Daunorubicin 60 mg/sqm on days 1-3) + Gemtuzumab Ozogamicin intravenously 3 mg/m² (dose capping at 5 mg] on days 1, 4, and 7. Patients in CR received two consolidation courses of intravenous daunorubicin (60 mg/m² for 1 day or 2 days) in combination with intravenous ARA-C (1000 mg/sqm iv bid on days 1–4) + Gemtuzumab Ozogamicin 3 mg/m² (dose capping at 5 mg] on days 1 [reference: Castaigne S, et al; Lancet. 2012; 379: 1508]. Patients with CBF-related AML were not allocated to allogeneic HSCT in first CR.

## S2. Multiparametric flow cytometry (MFC) methods for detection of aberrant Leukemia-Associate Immuno-Phenotypes (LAIP).

MFC study files reporting individual leukemia-aberrant immune-phenotype (LAIP) profiles were acquired locally according to pre-defined standard operating procedures. The same LAIP quantification was applied to BM samples for MRD assessment after induction and consolidation cycles. This evaluation was carried out at hematopoietic recovery and within 28 days after the end of chemotherapy in any instance. Acquisition through an SSC-antigen live-gate was performed and at least 8 x $10^5$ BM nucleated cells were collected. LAIP profiles for measurable residual disease (MRD) study were detected using multiple combinations including CD45 conjugated with peridinin chlorophyll protein (PerCP or PerCP-Cy5.5). The panel of diagnostic monoclonal antibodies (MoAb) was previously established and reported elsewhere[a]. A FACSCanto II flow cytometer (Becton Dickinson, BD, San Jose, CA) was used equipped with FACSDiva Software (BD) for data

acquisition. Instrument setup, calibration and quality control were performed to ensure measures' stability[b]. Consistency of fluorescence intensity was monitored weekly by running fluorochrome-conjugated beads (CS&T, BD). Fluorescence photomultiplier voltages were adjusted until the mean channel values for the unlabelled beads corresponded to predetermined target values. Overtime stability of bead mean fluorescence intensity (MFI) profile was checked by Levey-Jennings diagrams; changes of up to ±15% of the mean target MFI were tolerated. The mixed-bead suspension was used to determine the appropriate compensation settings. Each combination of MoAbs was added to 50 µl of a suspension of BM cells adjusted to 20,000 nucleated cells/µl; a stain-lyse-and-then-wash procedure was adopted.

**References**


a) Mannelli F, Gianfaldoni G, Bencini S, et al. Early peripheral blast cell clearance predicts minimal residual disease status and refines disease prognosis in acute myeloid leukemia. Am J Hematol 2020;95(11):1304–1313.

b) Owens MA, Vall HG, Hurley AA, Wormsley SB. Validation and quality control of immunophenotyping in clinical flow cytometry. Journal of Immunological Methods. 2000;243(1–2):33–50.

## S3. PCR-based MRD

Sensitive Real-time quantitative-polymerase chain reaction assays (RQ-PCR) was used for detection of MRD in patients with a suitable molecular probe. RQ-PCR was performed following the Europe Against Cancer (EAC) program recommendations[a] with a sensitivity of $10^{-5}$. Level of *RUNX1-RUNX1T1* and *CBFB-MYH11* transcripts and *NPM1* gene mutations were detected by Ipsogen commercial kits: Ipsogen *RUNX1-RUNX1T1* Kit, Ipsogen *CBFB-MYH11* A Kit, Ipsogen *NPM1* mutA MutaQuant Kit and Ipsogen *NPM1* mut B&D MutaQuant Kit (Qiagen, Courtaboeuf, France). One microgram of RNA was reverse transcribed according to EAC protocol. RQ-PCR was performed according to the manufacturer's instructions on a 7900 ABI platform (Applied Biosystems, Foster City, USA). Amplification conditions were: 2 min at 50 °C, 10 min at 95 °C followed by 50 cycles at 95 °C for 15 s and at 6 °C for 1 min. The *NPM1* mutations, *RUNX1-RUNX1T1* and *CBFB-MYH11* transcript values were normalized on the number of housekeeping gene Abelson (*ABL*) transcripts and were expressed as the number of target gene copies per $10^4$ copies of *ABL*. Using standards with a known number of molecules, it was possible to establish a standard curve and determine the precise amount of target in the test sample. The Ipsogen standard curves are plasmid-based: 3 plasmid standard dilutions for the control gene, and 5 standard dilutions for the mutated gene, to ensure accurate standard curves. All samples were analyzed in duplicate. A threshold value of 0.1 was used and baseline was set to 3–15 either for *ABL* or target genes[b].

**References**


a) Gabert J, Beillard E, Velden VHJ van der, 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.

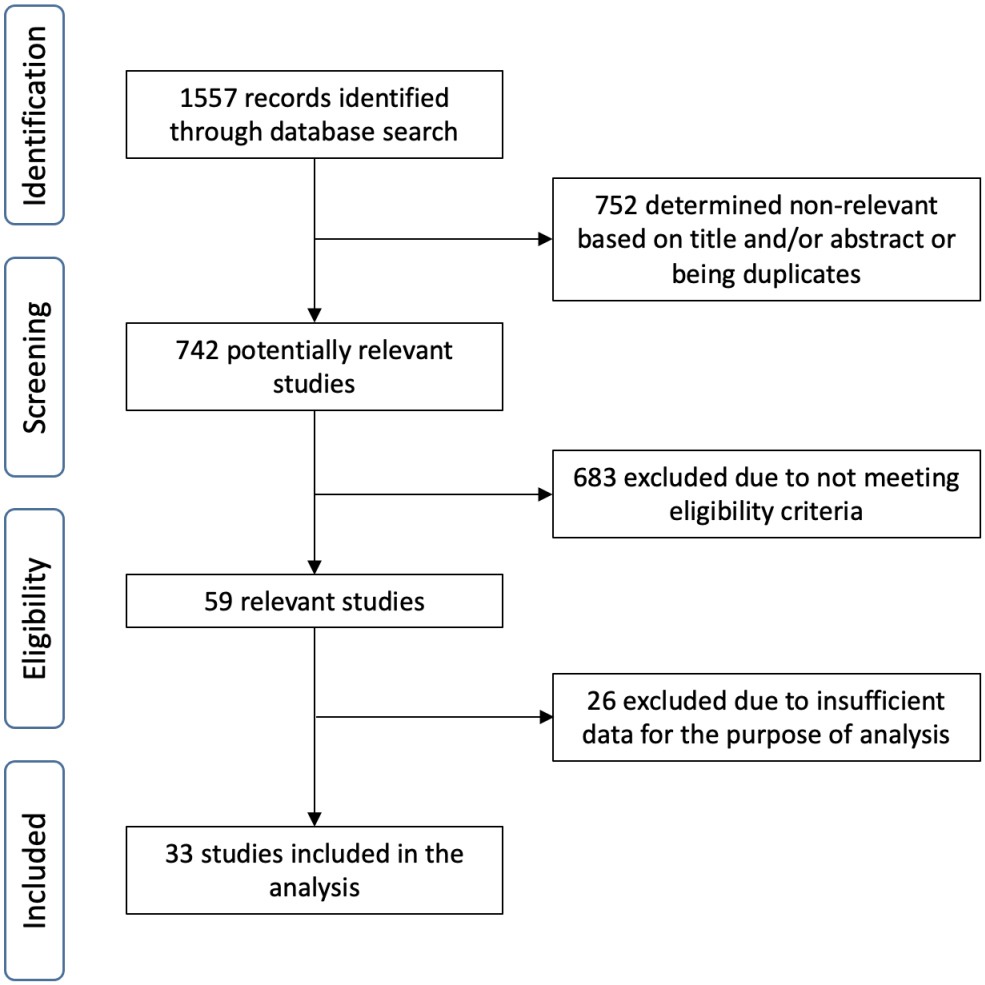
b) Gorello P, Cazzaniga G, Alberti F, et al. Quantitative assessment of minimal residual disease in acute myeloid leukemia carrying nucleophosmin (NPM1) gene mutations. Leukemia 2006;20(6):1103–1108.

## S4. Analysis of literature: description of methods and flow diagram of the study selection process

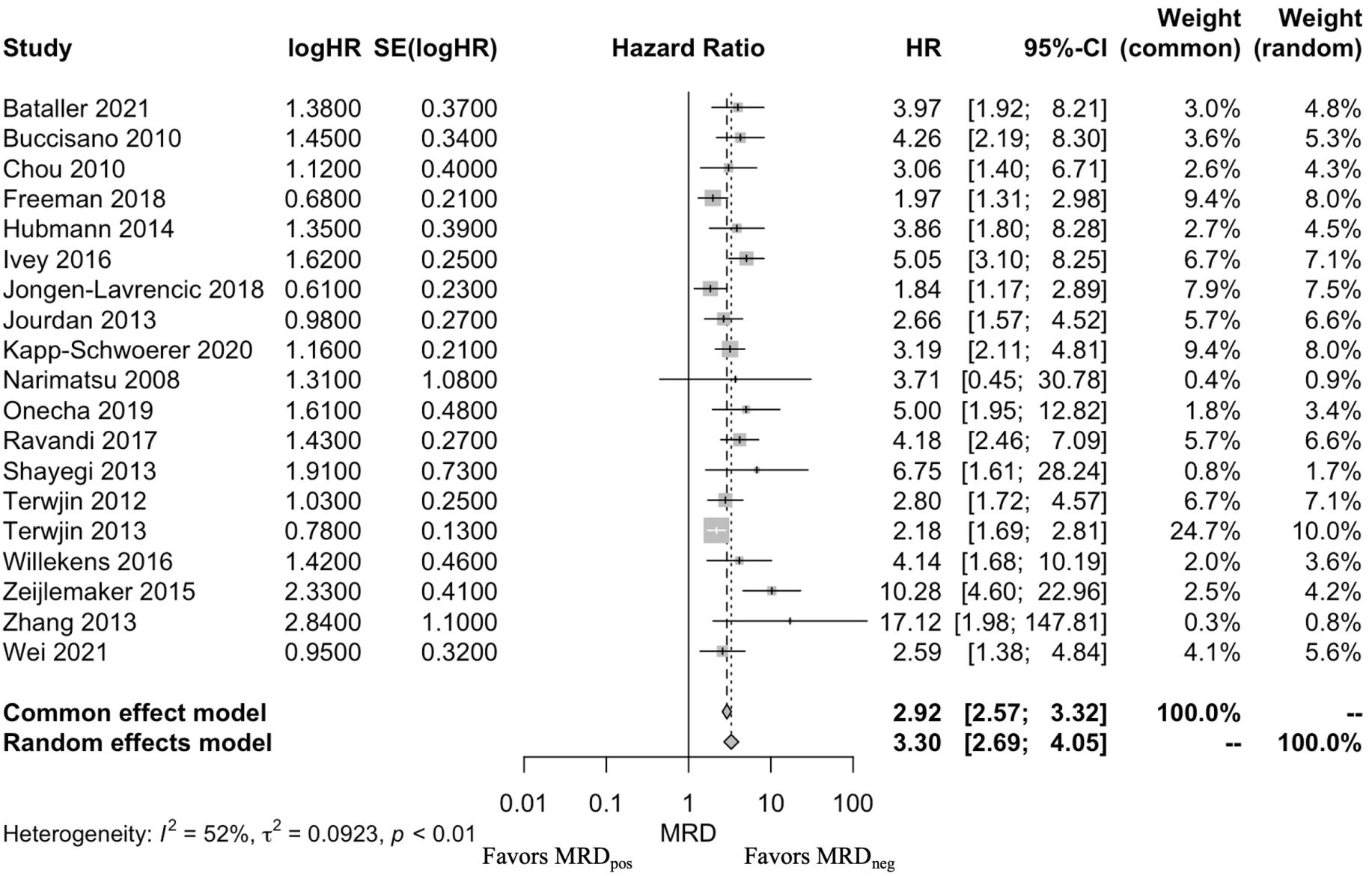
We carried out a search in PubMed for articles published between 2000 and 2021 by filtering for keywords (AML, acute myeloid leukemia, or acute myelogenous leukemia, and MRD, minimal residual disease, or measurable residual disease). The results are summarized in the Figures below. Reports were screened and filtered according to the following criteria: sufficiently detailed MRD data, sufficiently detailed treatment information, availability of Kaplan-Meier curves for DFS in the paper. Based on this assessment, a total of 33 articles were selected. The extracted data are detailed in Table S1. We extracted survival data from Kaplan-Meier curves by using the commercial graph digitizer software (Digitizelt, version 2.1; Bormisoft) and applying a previously published algorithm to reconstruct survival data for $MRD_{pos}$ and $MRD_{neg}$ cases[a]. The main characteristics of treatment in the first two cycles (drugs, ARA-C dosage, schedule) were obtained for each report and tabulated as in Supplemental Table S2. Moreover, each extracted case was annotated for the following variables: genetic subset, method for MRD detection, number of chemotherapy cycles pre-MRD assessment, cumulative dosage of ARA-C pre-MRD assessment, chemotherapy schedule pre-MRD, MRD status. In case of multiple MRD time-points, results were extrapolated and annotated accordingly. Studies selected for analysis of DFS in $MRD2_{neg}$ cases based on treatment intensity were processed as in conventional meta-analyses and extracted data are summarized in a Forest plot (see below S5).

**References**


a) Guyot P, Ades A, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. Bmc Med Res Methodol 2012;12(1):9


Identification
Screening
Eligibility
Included
1557 records identified through database search
752 determined non-relevant based on title and/or abstract or being duplicates
742 potentially relevant studies
683 excluded due to not meeting eligibility criteria
59 relevant studies
26 excluded due to insufficient data for the purpose of analysis
33 studies included in the analysis

**S5. Forest plot summarizing the effects of MRD as assessed by hazard ratio (HR), standard error (SE), and the relative weight of each study included in the analysis of $MRD_{neg}$ patients according to the intensity of treatment.**

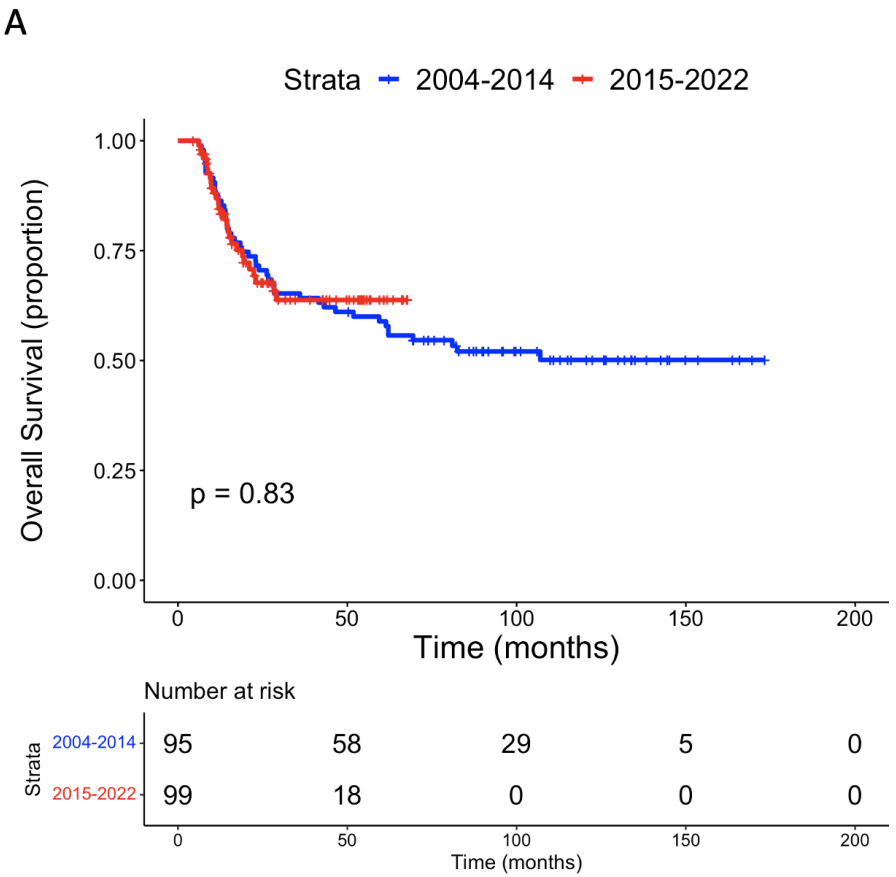


**S6. Statistical methods**

Pairwise comparisons of patient characteristics between groups, as defined by MRD, baseline features and treatment intensity, were performed using the Mann-Whitney test or the Kruskal-Wallis test for continuous variables and Pearson's chi-squared test or Fisher's exact test for categorical variables. Survival was estimated with the Kaplan-Meier method and long-term outcomes were compared with the log-rank test. The Cox proportional-hazards model was applied to estimate hazard ratios with 95% confidence intervals (CI) for DFS (the interval from CR to relapse or death), and OS (the interval from study entry to death) in both univariate and multivariate contexts. Comparison among longitudinal MRD assessments was done through the Harrells' concordance index (C-index) and 95% CIs, to evaluate the ability of the individual MRD time-point to predict outcome. To rule out an impact by allogeneic SCT, we censored patients receiving allogeneic SCT at the date of transplant in a further analysis. All P values were two-sided, and a 5% significance level was set.

## Supplemental Figures and Legends

**Figure S1.** Overall survival according to year of diagnosis, separating the patient series in two (A), and three (B) consecutive time periods.



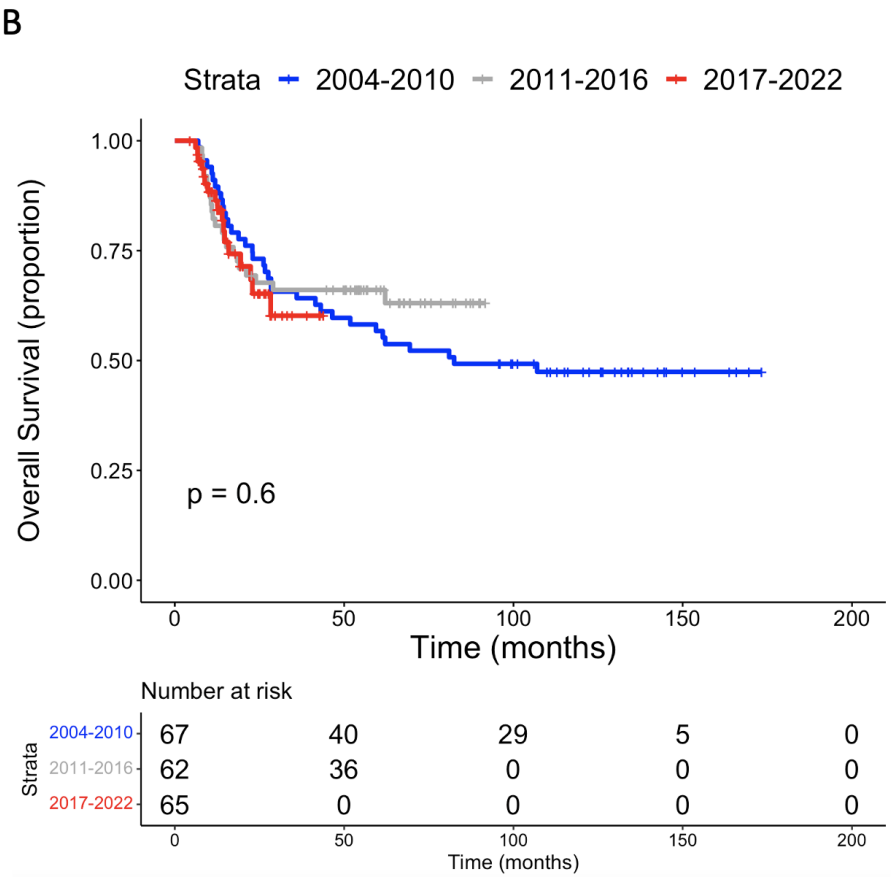



**Figure S2.** Disease-free (A), and overall (B) survival according to MRD1 status in the overall cohort.

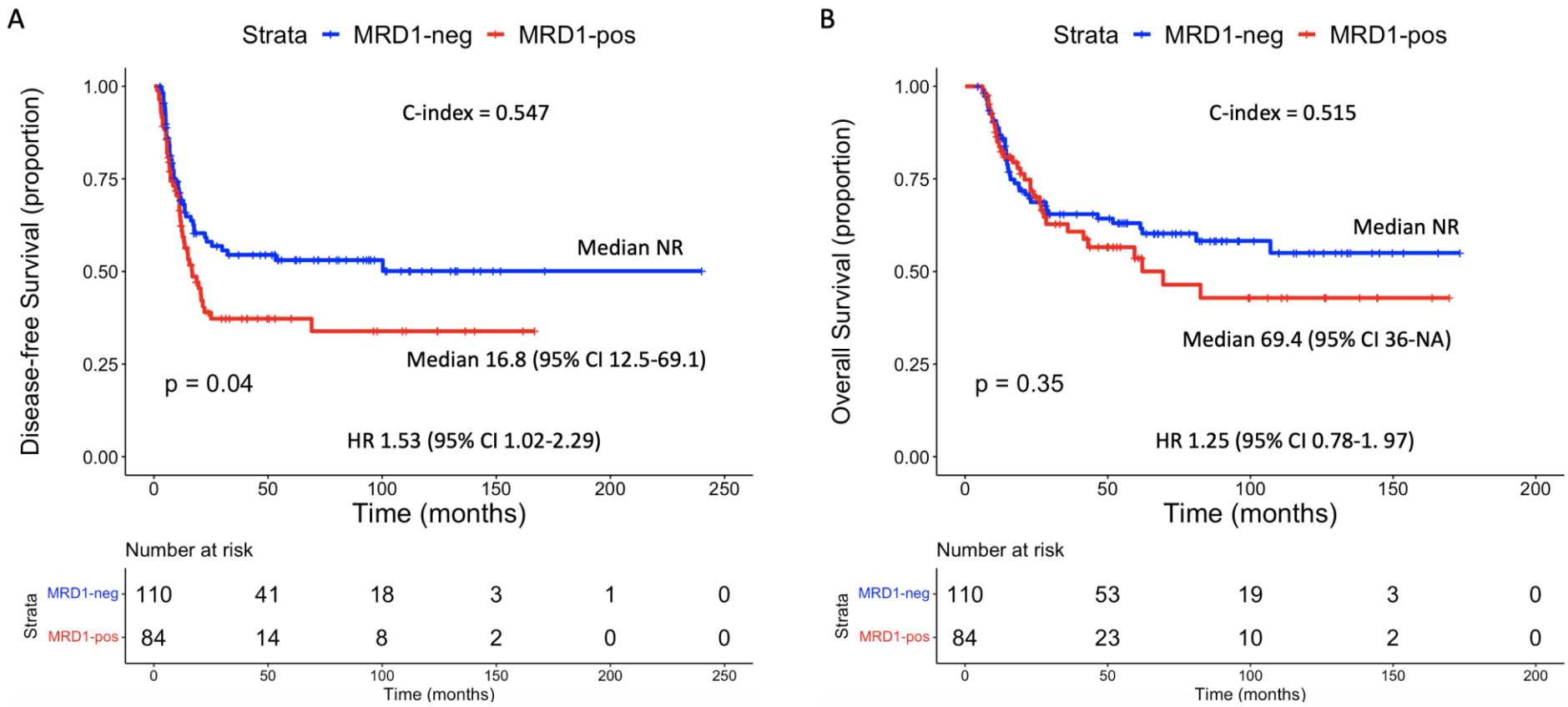


**Figure S3.** Disease-free (A), and overall (B) survival according to MRD2 status in the overall cohort.

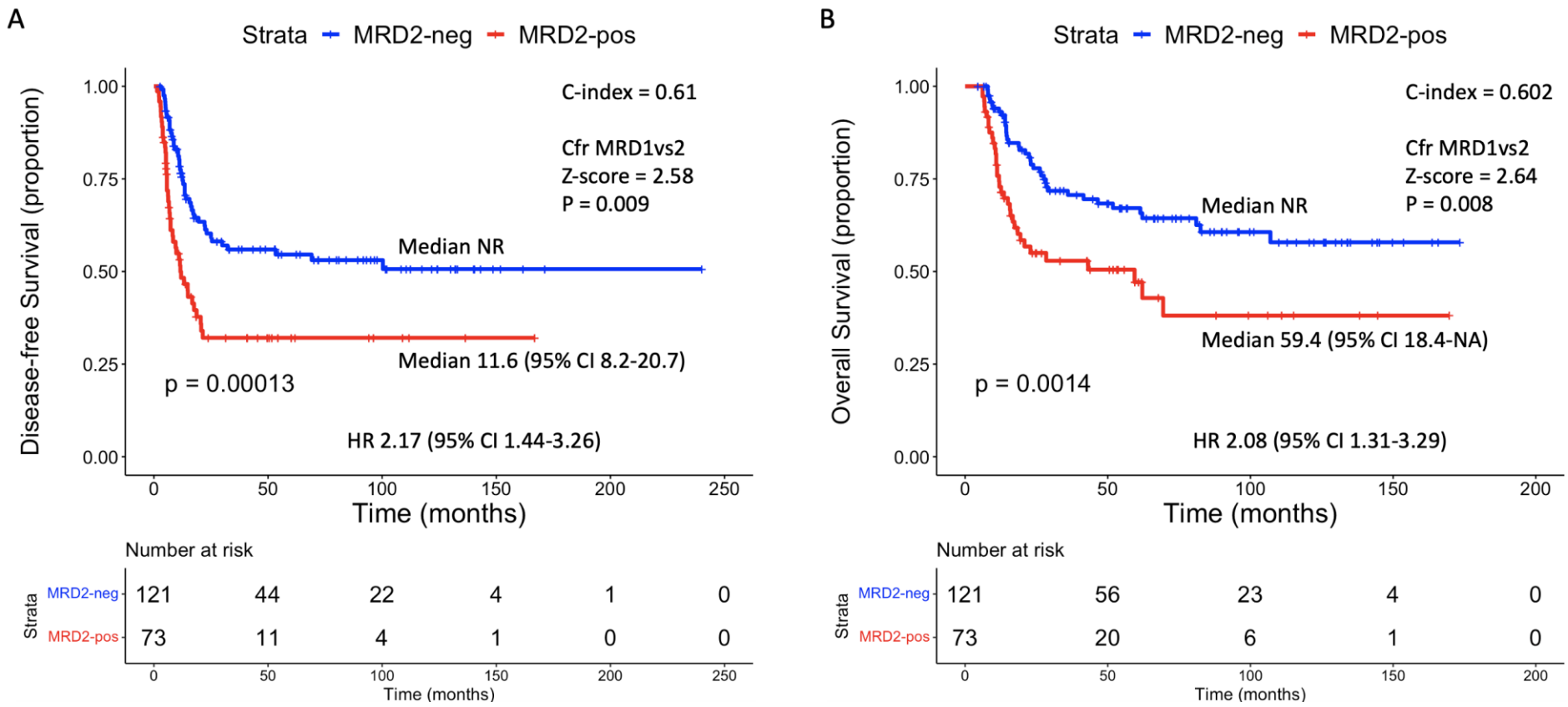


**Figure S4.** Disease-free (A-C), and overall (B-D) survival according to MRD2 status in age-related strata: patients aged < 55y (A-B) and ≥ 55y (C-D).

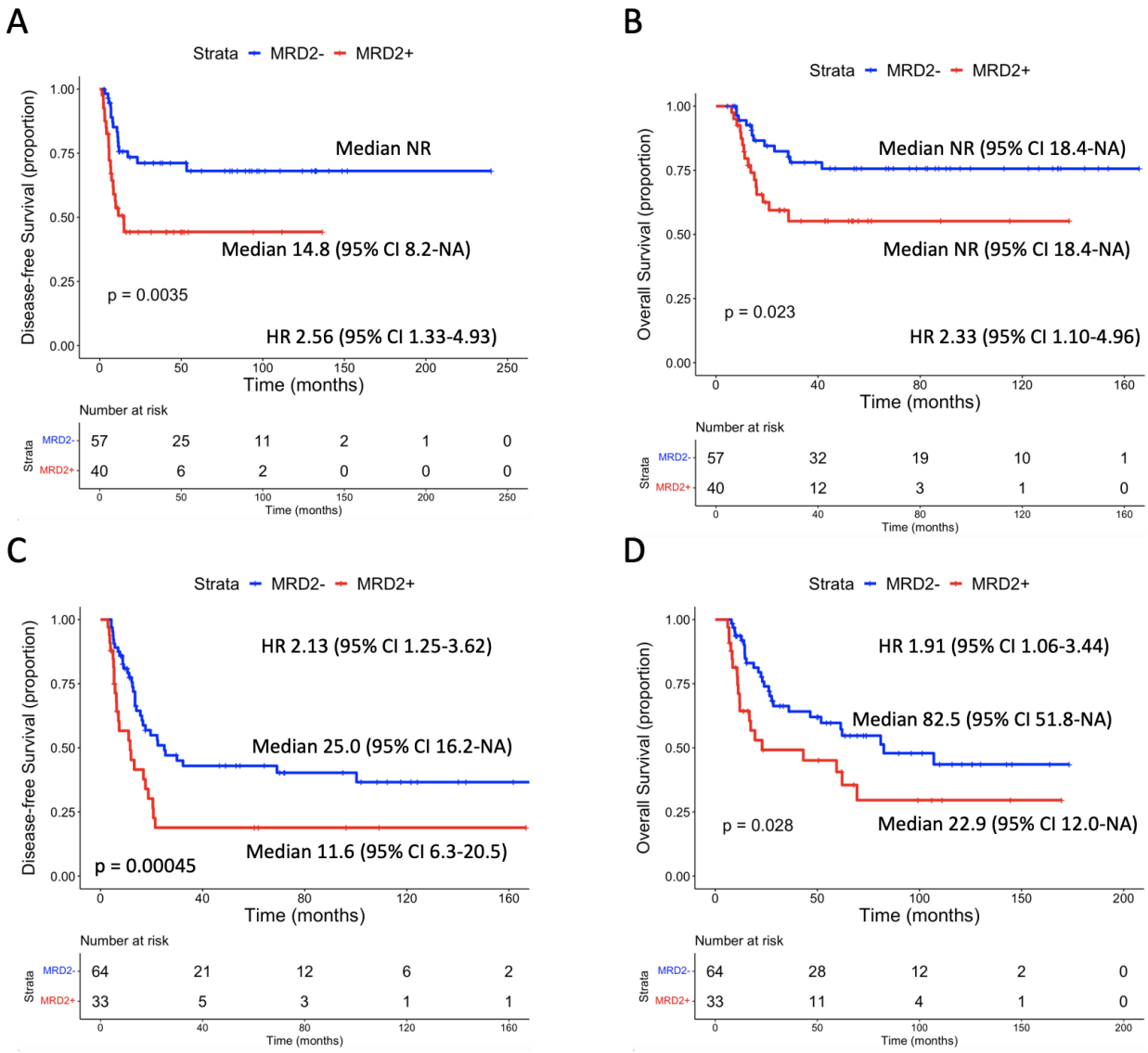


**Figure S5.** Disease-free (A-C), and overall (B-D) survival according to MRD2 status in WBC-related strata: WBC < $30x10^9$/L (A-B) and ≥ $30x10^9$/L (C-D).

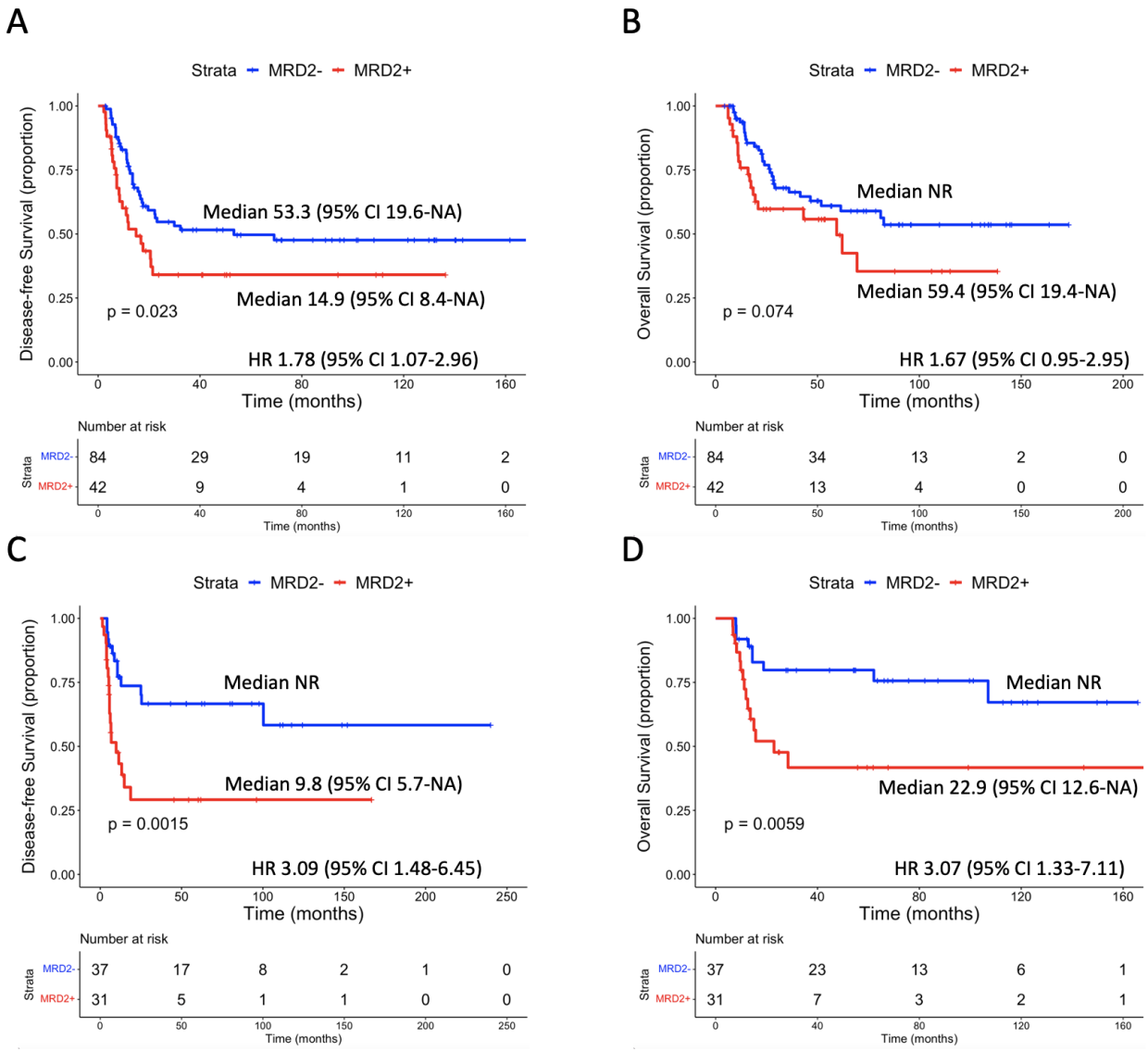


**Figure S6** Disease-free (A-C), and overall (B-D) survival according to MRD2 status in ELN-related strata: ELN favorable (A-B) and ELN intermediate (C-D).

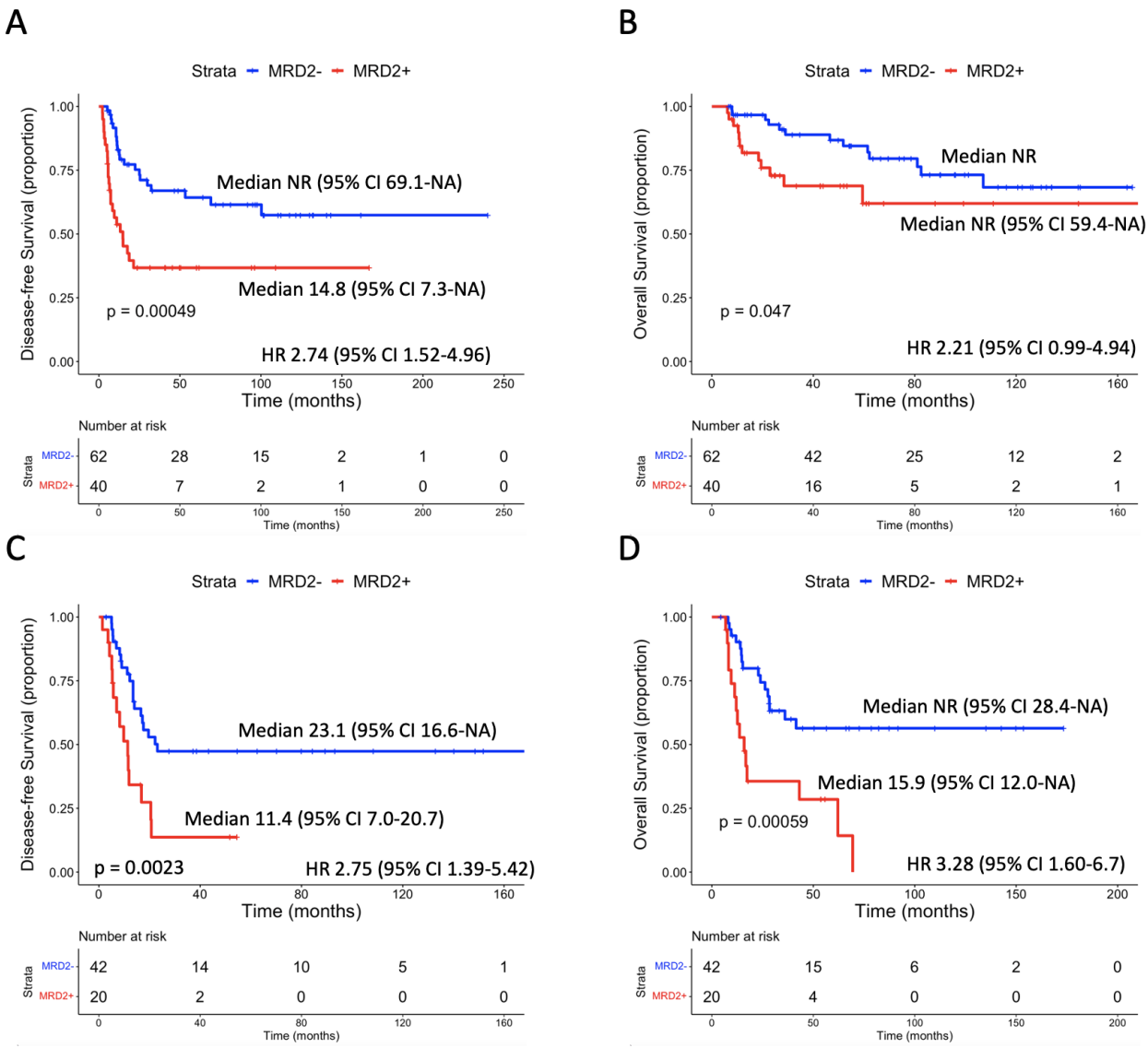


**Figure S7.** Disease-free (A-C), and overall (B-D) survival according to MRD2 status in gender-related strata: female (A-B) and male (C-D) patients.

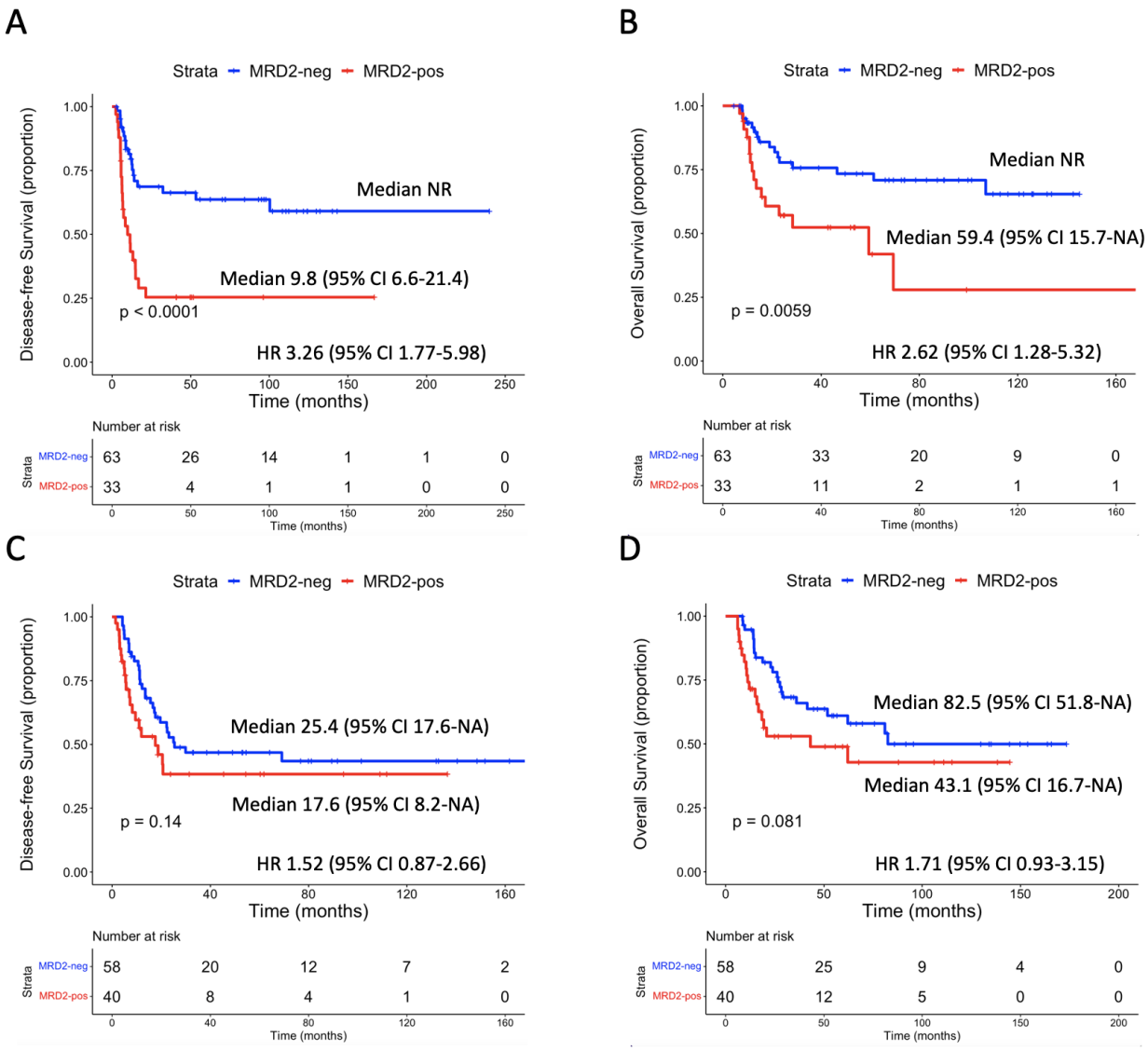


**Figure S8.** Disease-free survival of $MRD2_{neg}$ patients according to age- (A), WBC- (B), ELN- (C) and gender- (D) related categories.

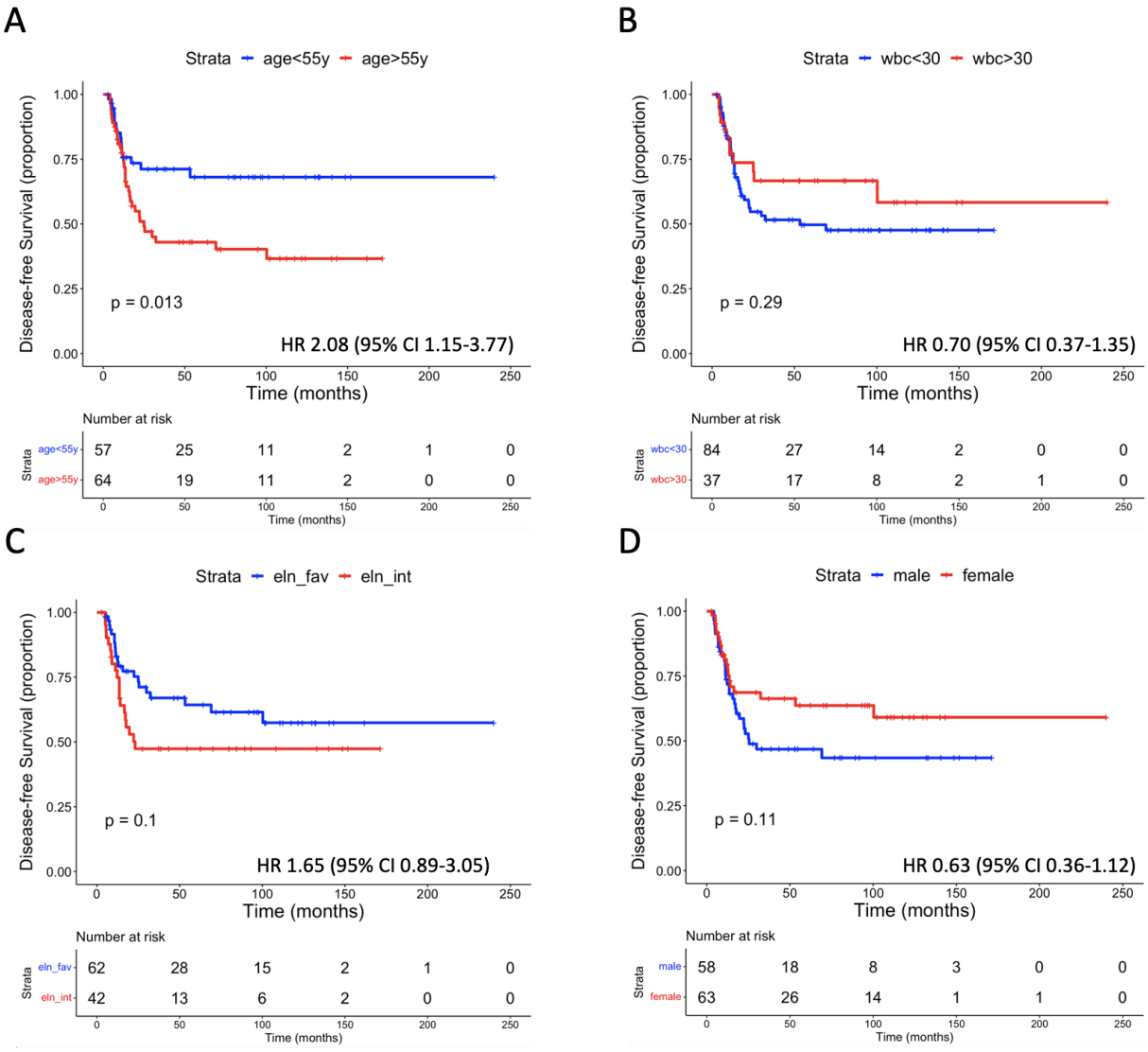


**Figure S9.** Disease-free (A-C) and overall (B-D) survival according to treatment intensity in first induction cycle (A-B) and in induction + first consolidation cycles (C-D).

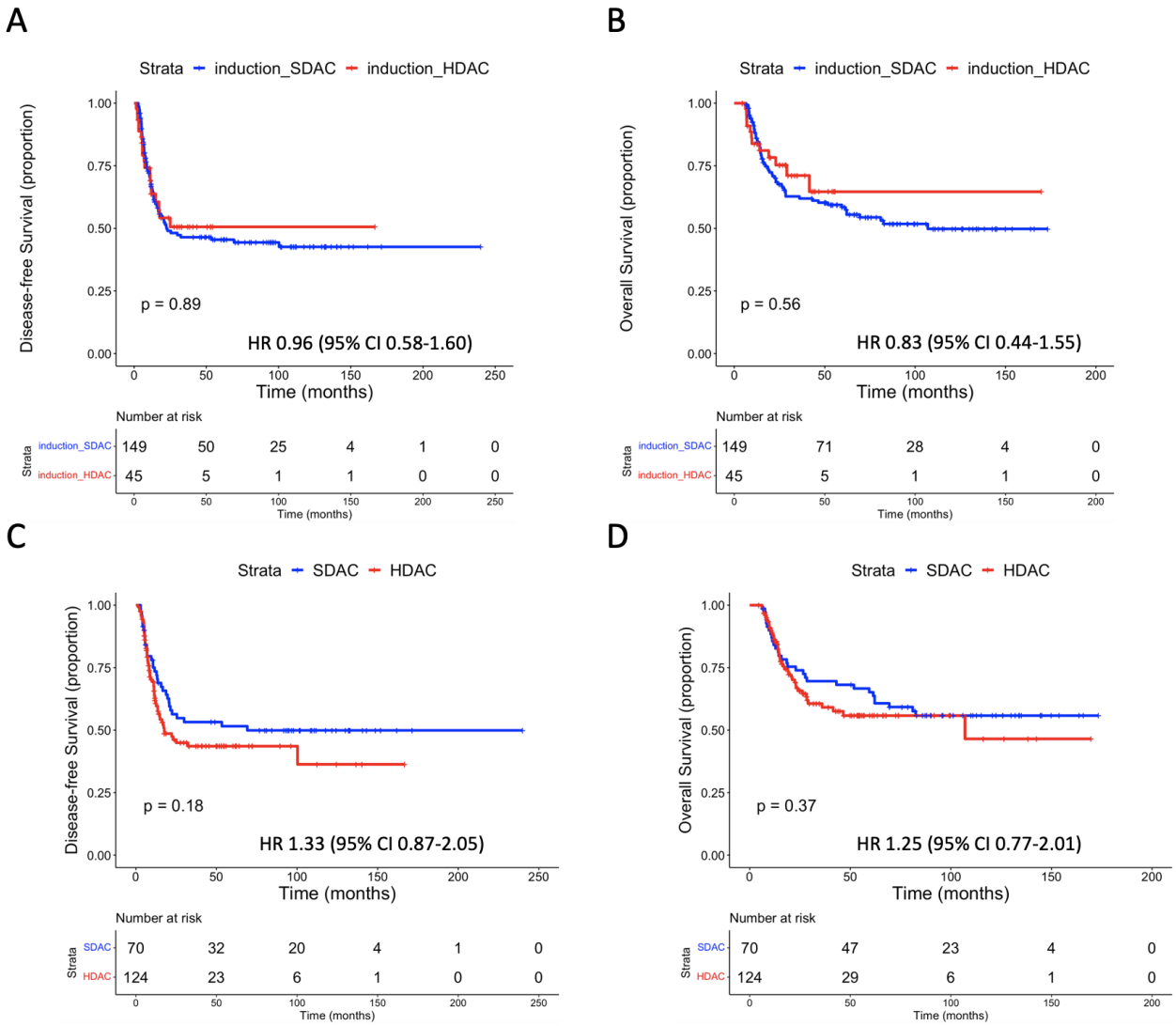


**Figure S10.** Disease-free (A-C) and overall (B-D) survival according to MRD2 status in treatment intensity categories within intermediate-risk karyotype: standard dose cytarabine (SDAC, panels A-B) and high-dose cytarabine (HDAC, panels C-D).

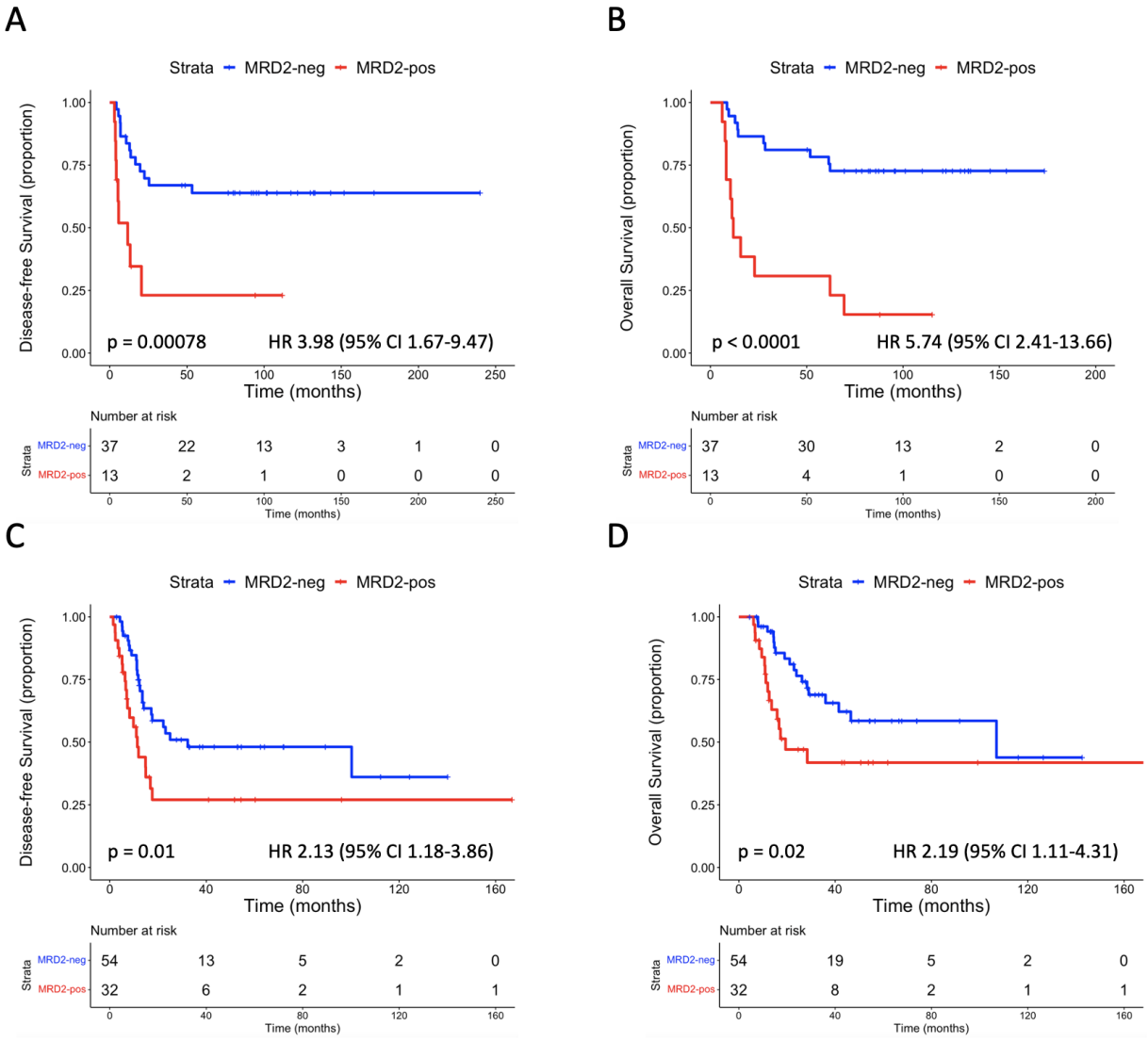


**Figure S11.** Disease-free (A) and overall (B) survival in $MRD2_{neg}$ patients according to treatment intensity.

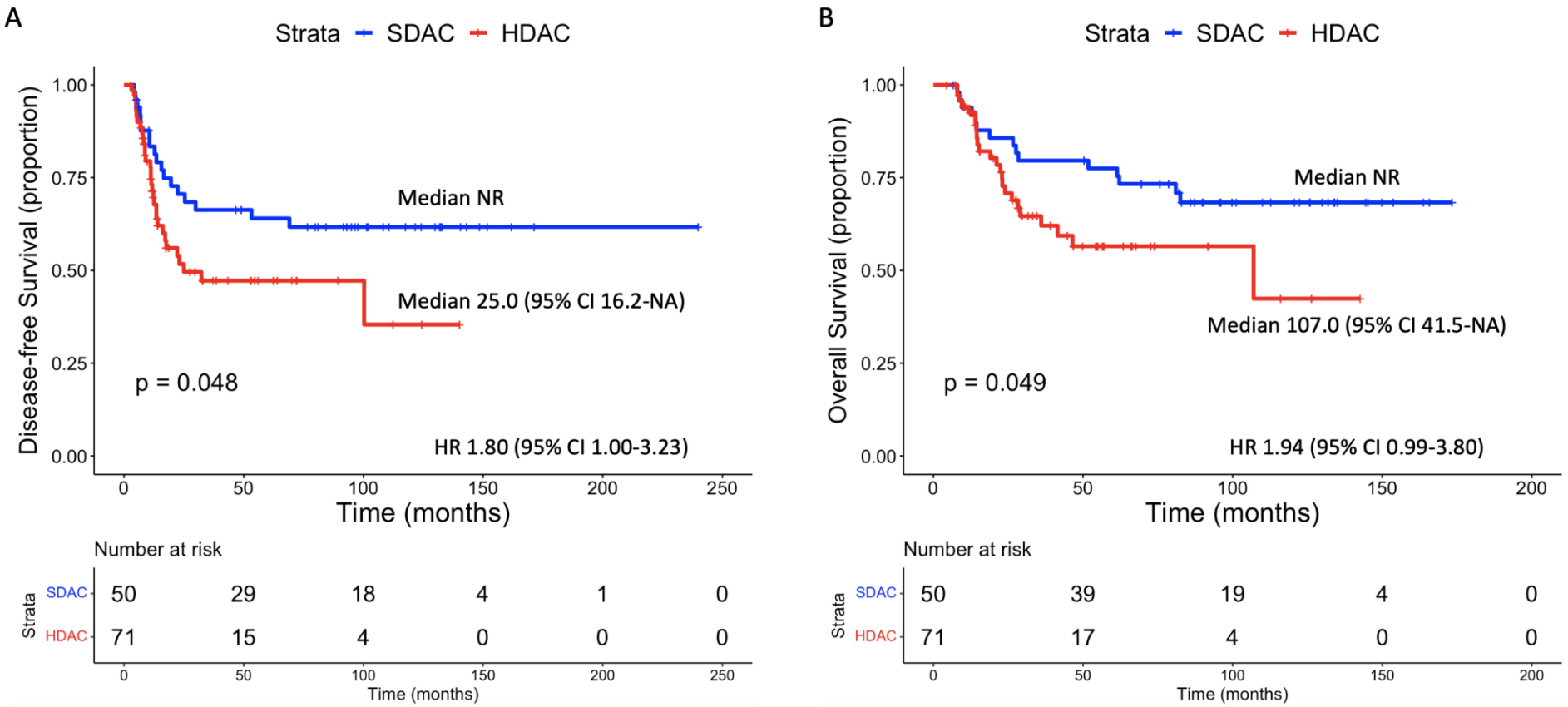


**Figure S12.** Disease-free (A-C) and overall (B-D) survival in $MRD2_{neg}$ patients according to treatment intensity categories within intermediate-risk karyotype (A-B) and ELN 2017 (C-D) categories.

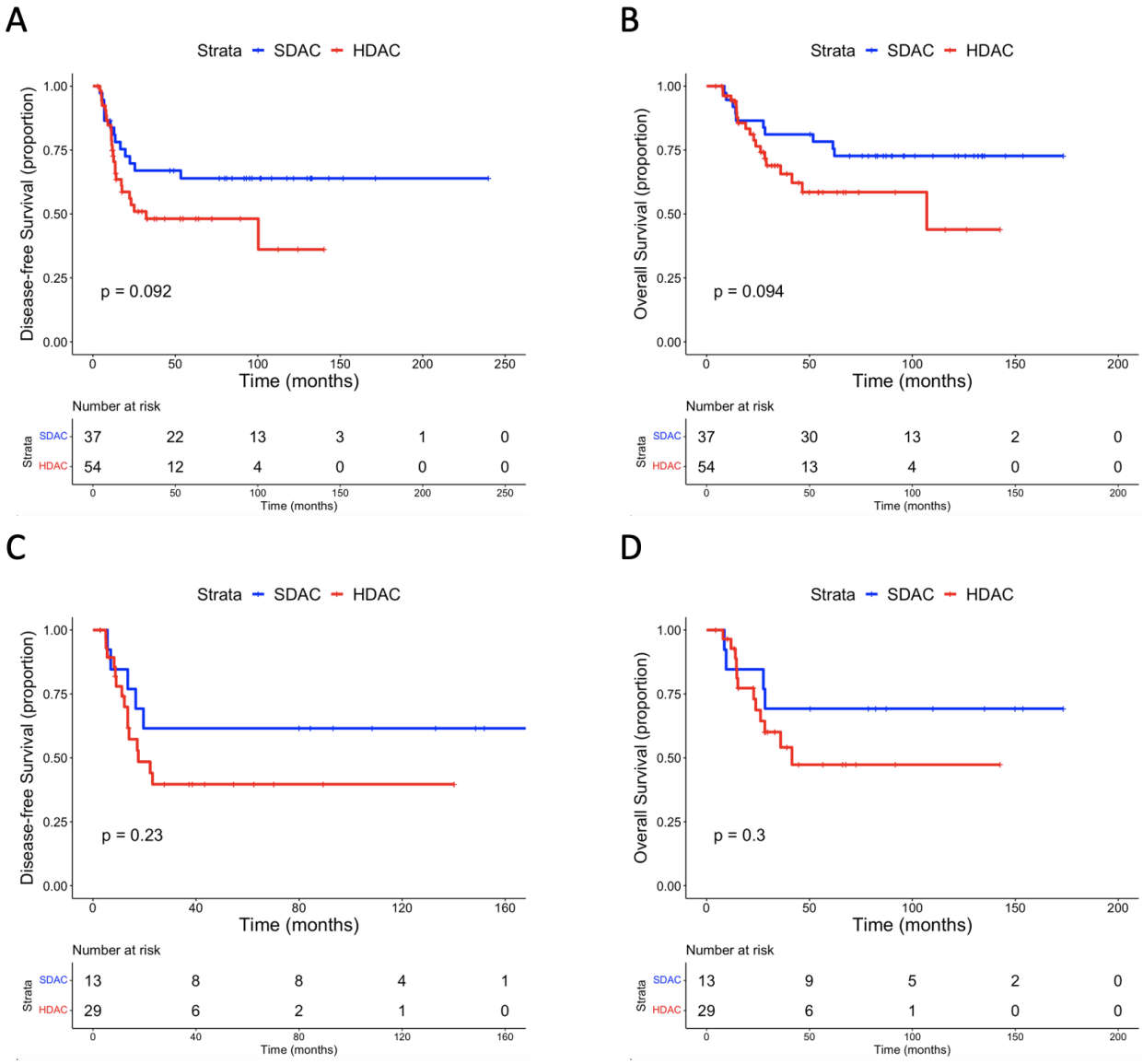


**Figure S13.** Disease-free (A-C) and overall (B-D) survival in $MRD2_{neg}$ patients according to combined model (age and treatment intensity) within intermediate-risk karyotype (A-B) and ELN (C-D) categories.

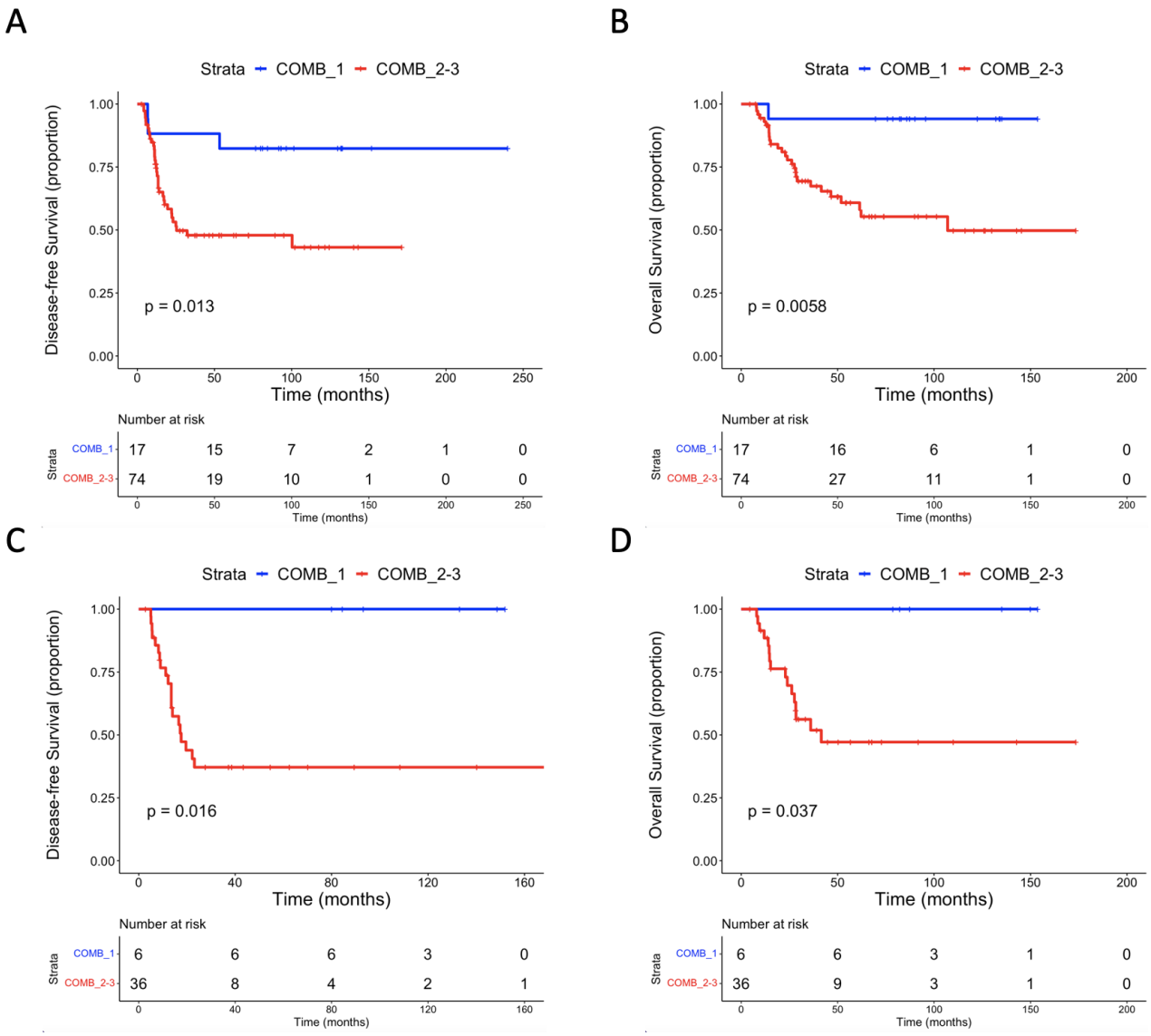
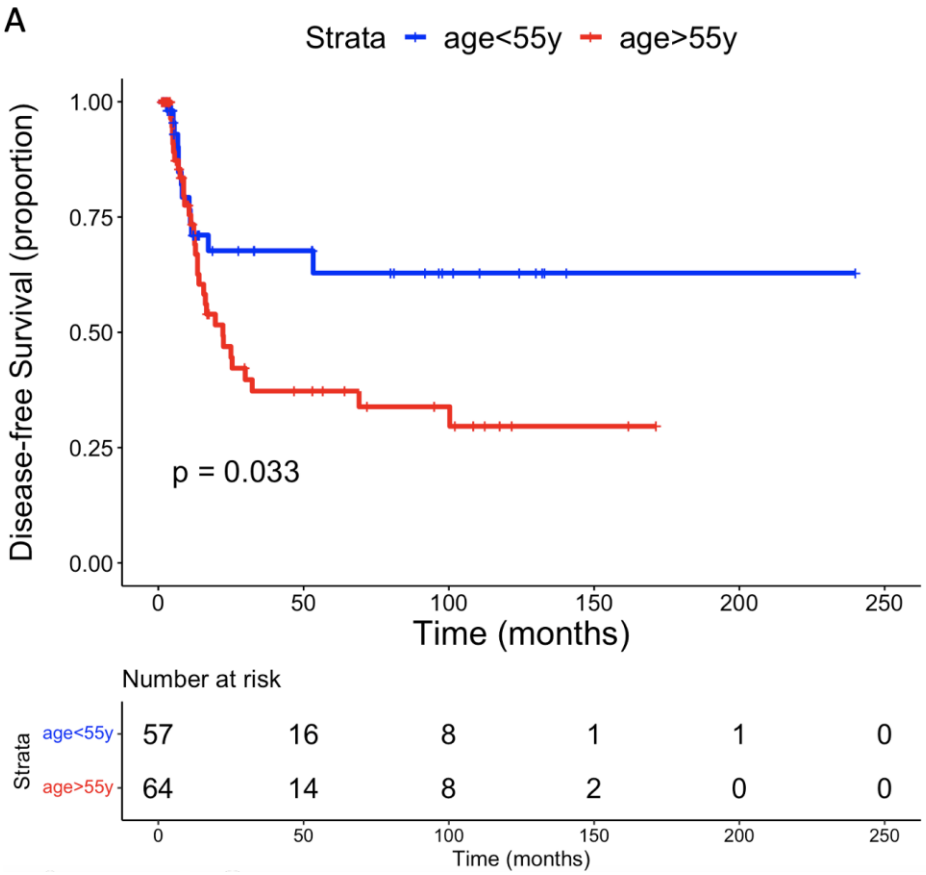


**Figure S14.** Disease-free survival in $MRD2_{neg}$ patients according to age (A) and treatment (B) categories after censoring at allogeneic transplant.



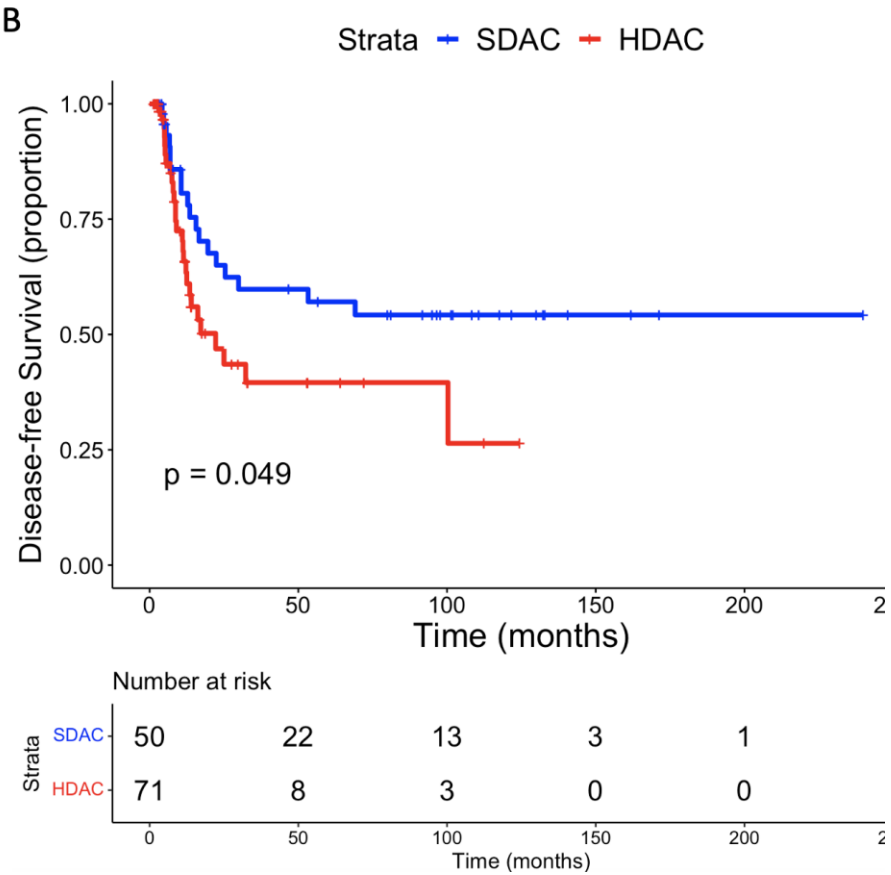



**Figure S15** Disease-free (A) and overall (B) survival in $MRD2_{neg}$ patients according to combined model category after censoring at allogeneic transplant: younger, SDAC treated (COMB_1) versus elderly and/or HDAC-treated (COMB_2-3) patients.

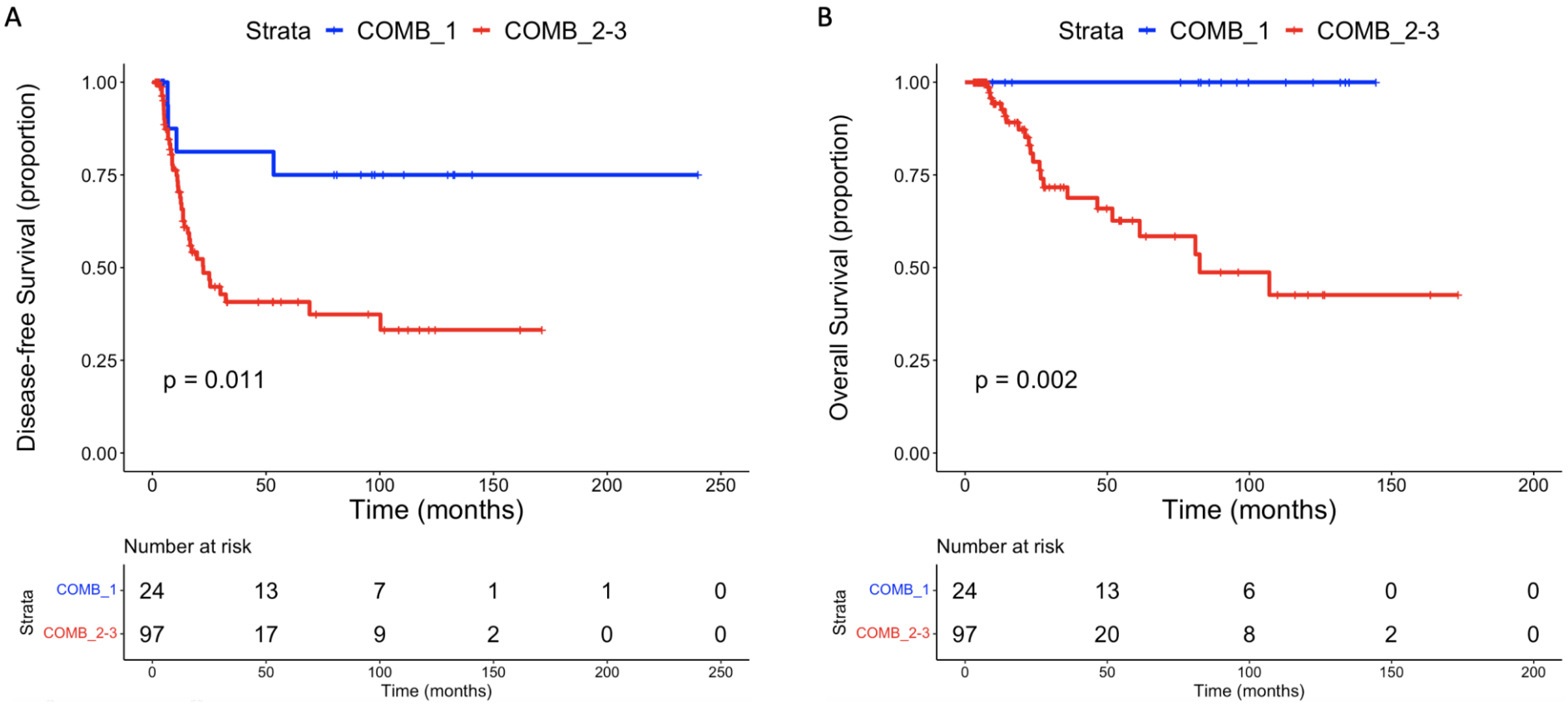
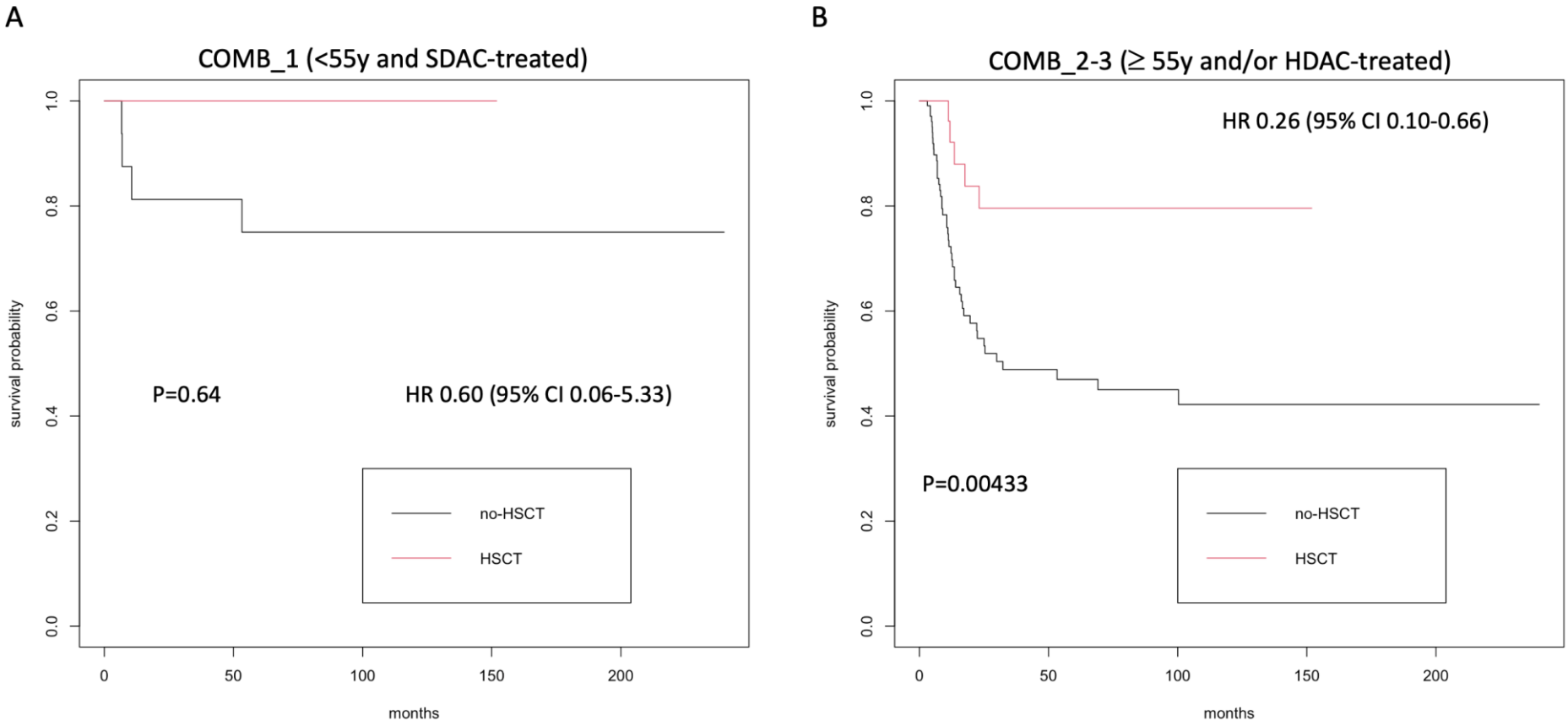


**Figure S16.** Effect of allogeneic HSCT on disease-free survival as depicted by Simon-Makuch plots in younger, SDAC-treated (A) and elderly and/or HDAC-treated (B) patients.



## Supplemental Tables and Results

**Table S1.** Analysis of literature: summary of the selected clinical trials

| Reference | Pts | Subset | Median age (range) | Method | Threshold for MRD status definition | Time-point |
|---|---|---|---|---|---|---|
| **Balsat 2017** | 152 | *NPM1*-mut | 49 (21-61) | RT-PCR | 4-log reduction | Post induction |
| **Bataller 2021** | 110 | *NPM1*-mut | 54 (18-71) | RT-PCR | Any positivity; ratio on *ABL*1 | Post induction; post consolidation |
| **Boddu 2018** | 104 | *CBF*-AML | 53-19-81 | RT-PCR | Slope of log-reduction | Post induction |
| **Buccisano 2010** | 143 | Unselected | NA (72% <60 y) | MFC | 0.035% of total cells | Post consolidation |
| **Chou 2010** | 55 | *FLT3*-ITD | 49 (17-90) | RT-PCR | 3-log reduction | Post consolidation |
| **Corbacioglu 2010** | 84 | *CBF::MYH11* | NA (16-60) | RT-PCR | Transcript copies reduction; any positivity | Post consolidation |
| **Ferret 2018** | 103 | *IDH1/2*-mut | 54 (22-70) | ddPCR | Detection limit 0.2% | Post induction |
| **Frairia 2017** | 223 | Unselected | 56 (19-76) | RT-PCR | *WT1*; 2-log reduction | Post induction |
| **Freeman 2018** | 286 | Unselected | 50 (16-71) | MFC | Variable between 0.05-0.2% based on controls | Post induction; post consolidation |
| **Gueieze 2010** | 59 | *CBF::MYH11* | 36 (4-77) | RT-PCR | 3-log reduction at MRD2; 0.001% at MRD3 | Post consolidation |
| **Hubmann 2014** | 158 | *NPM1*-mut | 57 (18-80) | RT-PCR | 0.01%; 3-log reduction | Post induction; post consolidation |
| **Ivey 2016** | 346 | *NPM1*-mut | 50 (6-68) | RT-PCR | Any positivity | Post consolidation |
| **Jongen-Lavrencic 2018** | 430 | Unselected | 51 (18-66) | NGS | Any positivity | Post consolidation |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Jourdan 2013** | 198 | *CBF*-AML | 42 (18-60) | RT-PCR | 3-log reduction | Post consolidation |
| **Kapp-Schwoerer 2020** | 469 | *NPM1*-mut | 58 (20-78) | RT-PCR | Any positivity | Post induction |
| **Klco 2015** | 68 | Unselected | 50 (39-58) | NGS | VAF 2.5% | Post induction |
| **Kronke 2011** | 245 | *NPM1*-mut | 49 (19-61) | RT-PCR | Any positivity | Post consolidation |
| **Marani 2013** | 42 | Unselected | 54 (17-81) | RT-PCR | *WT1*; 1.5-log reduction | Post induction |
| **Morita 2018** | 131 | Unselected | 51 (NA) | NGS | VAF strata | Post induction |
| **Narimatsu 2008** | 46 | *RUNX1::RUNXT1* | 50 (25-64) | RT-PCR | 1000 copies | Post consolidation |
| **Onecha 2019** | 63 | *NPM1 - IDH1/2-FLT3*-mut | 54 (NA) | NGS, RT-PCR, MFC | Dependent on method and time-point | Post induction; post consolidation |
| **Othus 2016** | 170 | Unselected | NA (18-60) | MFC | 0.01% of total cells | Post induction |
| **Ravandi 2017** | 186 | Unselected | 51 (17-79) | MFC | 0.1-0.01% of total cells | Post induction; post consolidation |
| **Rossi 2014** | 45 | Unselected | 53 (19-76) | MFC, RT-PCR | Dependent on method; log-reduction and clearance | Post induction; post consolidation |
| **Shayegi 2013** | 92 | *NPM1*-mut | 51 (20-79) | RT-PCR | Strata (negative, 0.1-1%, >1%) | Post induction; post consolidation |
| **Terwjin 2012** | 77 | Unselected | NA (NA) | MFC | Dependent on time-point | Post induction; post consolidation |
| **Terwjin 2013** | 164 | Unselected | 48 (18-60) | MFC | Dependent on time-point | Post induction; post consolidation |
| **Willekens 2016** | 94 | *RUNX1::RUNXT1* | 41 (18-60) | RT-PCR | Ratio on *ABL1* >0.001% | Post consolidation |
| **Yin 2012** | 278 | *CBF*-AML | 42 (15-70) | RT-PCR | Log-reduction for *RUNX1::RUNXT1;* number of copies for *CBF::MYH11* | Post induction; post consolidation |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Yoon 2014** | 206 | *CBF*-AML | 39 (18-89) | RT-PCR | 3 log-reduction at MRD1; number of copies at MRD2 | Post induction; post consolidation |
| **Zeijlemaker 2015** | 114 | Unselected | 59 (25-73) | MFC | Dependent on LAIP and sample type | Post induction; post consolidation |
| **Zhang 2013** | 52 | *RUNX1::RUNXT1* | 21 (13-57) | RT-PCR | Ratio on *ABL1* >0.01% | Post consolidation |
| **Wei 2021** | 187 | *RUNX1::RUNXT1* | 34 (14-54) | RT-PCR | 3 log-reduction | Post consolidation |

**Table S2.** Analysis of literature: treatment details of the selected clinical trials.

| **Reference** | **Induction** | | | | | **Consolidation** | | |
|---|---|---|---|---|---|---|---|---|
| | **Trial** | **Scheme** | **ARA-C cumulative dose, mg** | **Anthracycline, cumulative dose, mg** | **Third drug, cumulative dose, mg** | **ARA-C dose, mg** | **Anthracycline, dose, mg** | **Third drug, dose, mg** |
| **Balsat 2017** | ALFA 07-02 | - | 7.500 | Daunorubicin 285 | - | 18.000 | - | - |
| **Bataller 2021** | - | - | 1.400 | Idarubicin 36 | - | 18.000 | - | - |
| **Boddu 2018** | - | - | 10.000 | Idarubicin 36 | Gemtuzumab 3 | 8.000 | - | Gemtuzumab 3 |
| | | | 8.000 | Idarubicin 36 | - | 10.000 | - | - |
| **Buccisano 2010** | AML-10 | DAE | 1.000 | Daunorubicin 150 | Etoposide 500 | 6.000 | Daunorubicin 150 | - |
| | AML-10 | ICE | 1.000 | Idarubicin 30 | - | 6.000 | Idarubicin 30 | - |
| | AML-10 | MiCE | 1.000 | Mitoxantrone 36 | - | 6.000 | Mitoxantrone 36 | - |
| | AML-12 | HDAC | 24.000 | Daunorubicin 150 | Etoposide 250 | 6.000 | Daunorubicin 150 | - |
| | AML-13 | MiCE | 700 | Mitoxantrone 21 | Etoposide 250 | 500 | Idarubicin 24 | - |
| | AML-17 | MiCE | 700 | Mitoxantrone 21 | Etoposide 300 | 500 | Idarubicin 24 | Etoposide 300 |
| **Chou 2010** | - | 3+7 | 700 | Idarubicin 36 | - | 16.000 | - | - |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Corbacioglu 2010** | AML HD93 | Double induction (ICE-ICE) | 1.400 | Idarubicin 72 | Etoposide 600 | 18.000 | Mitoxantrone 36 | - |
| | AML HD98A | Double induction (ICE-ICE) | 1.400 | Idarubicin 72 | Etoposide 600 | 18.000 | Mitoxantrone 36 | - |
| | AMLSG 07-04 | Double induction (ICE-ICE) | 1.400 | Idarubicin 72 | Etoposide 600 | 18.000 | Mitoxantrone 36 | - |
| **Ferret 2018** | ALFA 07-01 | 3+7 | 1.400 | Daunorubicin 180 | - | 8.000 | Daunorubicin 60 | - |
| | ALFA 07-01 | 3+7+GO | 1.400 | Daunorubicin 180 | Gentuzumab 9 | 8.000 | Daunorubicin 60 | Gentuzumab 3 |
| | ALFA 07-02 | - | 7.500 | Daunorubicin 285 | - | 18.000 | - | - |
| **Frairia 2017** | NILG 02-06 | ICE | 1.400 | Idarubicin 36 | Etoposide 800 | 1.400 | Idarubicin 36 | - |
| | - | Ida-FLA | 4.000 | Idarubicin 24 | Fludarabine 100 | 4.000 | Idarubicin 24 | Fludarabine 100 |
| **Freeman 2018** | MRC AML-17 | DA | 2.000 | Daunorubicin 150 | Gemtuzumab 3-6 | 2.000 | Daunorubicin 150 | - |
| | MRC AML-17 | ADE | 2.000 | Daunorubicin 150 | Etoposide 500<br>Gemtuzumab 3-6 | 2.000 | Daunorubicin 150 | - |
| **Gueieze 2010** | ALFA | 3+7 | 1.400 | Daunorubicin 180 | - | 18.000 | - | - |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Hubmann 2014** | AMLCG | Double induction (TAD-HAM) | 19.400 | Daunorubicin 180<br>Mitoxantrone 30 | Thioguanine 1.400 | 1.400 | Daunorubicin 180 | Thioguanine 1.400 |
| | | Double induction (HAM-HAM) | 36.000 | Mitoxantrone 60 | Thioguanine 1.400 | 1.400 | Daunorubicin 180 | Thioguanine 1.400 |
| | | Sequential induction (S-HAM) | 24.000 | Mitoxantrone 40 | Thioguanine 1.400 | 1.400 | Daunorubicin 180 | Thioguanine 1.400 |
| **Ivey 2016** | MRC AML-17 | DA | 2.000 | Daunorubicin 150 | Gemtuzumab 3-6 | 2.000 | Daunorubicin 150 | - |
| | MRC AML-17 | ADE | 2.000 | Daunorubicin 150 | Etoposide 500<br>Gemtuzumab 3-6 | 2.000 | Daunorubicin 150 | - |
| **Jongen-Lavrencic 2018** | HOVON 42 | - | 1.400 | Idarubicin 36 | - | 12.000 | - | Amsacrine 360 |
| | HOVON 42 | - | 10.000 | Idarubicin 36 | - | 16.000 | - | Amsacrine 360 |
| | HOVON 102 | - | 1.400 | Idarubicin 36 | - | 12.000 | - | Amsacrine 360 |
| | HOVON 102 | - | 1.400 | Idarubicin 36 | Clofarabine 50 | 12.000 | - | Amsacrine 360<br>Clofarabine 50 |
| **Jourdan 2013** | CBF 2006 | 3+7 | 1.400 | Daunorubicin 180 | - | 18.000 | - | - |
| | CBF 2006 | - | 7.500 | Daunorubicin 285 | - | 18.000 | - | - |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Kapp-Schwoerer 2020** | AMLSG 09-09 | 3+7 | 1.400 | Idarubicin 36 | Etoposide 500 Gentuzumab 3 | 1.400 | Idarubicin 36 | Etoposide 500 Gentuzumab 36 |
| **Klco 2015** | - | 3+7 | 1.400 | Daunorubicin 180 | - | - | - | - |
| **Kronke 2011** | AML HD98A | Double induction (ICE-ICE) | 1.400 | Idarubicin 72 | Etoposide 600 | 18.000 | Mitoxantrone 36 | - |
| | AMLSG 07-04 | Double induction (ICE-ICE) | 1.400 | Idarubicin 72 | Etoposide 600 | 18.000 | Mitoxantrone 36 | - |
| **Marani 2013** | - | FLAI5 | 10.000 | Idarubicin 36 | Fludarabine 150 | - | - | - |
| **Morita 2018** | - | FIA | 5.000 | Idarubicin 30 | Fludarabine 150 | - | - | - |
| | - | CIA | 5.000 | Idarubicin 30 | Clofarabine 100 | - | - | - |
| | - | CLIA | 5.000 | Idarubicin 30 | Cladribine 25 | - | - | - |
| **Narimatsu 2008** | - | 3+7 | 700 | Idarubicin 36 | - | 18.000 | - | - |
| | - | 3+7 | 700 | Idarubicin 36 | - | 1.400 | - | - |
| **Onecha 2019** | - | 3+7 | 700 | Idarubicin 36 | - | 18.000 | - | - |
| **Othus 2016** | SWOG 0106 | - | 700 | Daunorubicin 135 | - | - | - | - |
| | SWOG 0106 | - | 700 | Daunorubicin 135 | Gemtuzumab 6 | - | - | - |
| **Ravandi 2017** | - | CIA | 5.000 | Idarubicin 30 | Clofarabine 100 | 4.000 | Idarubicin 20 | |
| | - | FIA | 5.000 | Idarubicin 30 | Fludarabine 150 | 4.000 | Idarubicin 20 | Fludarabine 60 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | - | FLAG-Ida | 10.000 | Idarubicin 30 | Fludarabine 150 | 8.000 | Idarubicin 20 | Fludarabine 60 |
| | - | CLIA | 5.000 | Idarubicin 30 | Cladribine 25 | 4.000 | Idarubicin 20 | - |
| **Rossi 2014** | AML 12 | - | 24.000 | Daunorubicin 150 | Etoposide 250 | 6.000 | Daunorubicin 150 | - |
| | - | FLAG | 10.000 | Daunorubicin 30 | Fludarabine 100 | 6.000 | Daunorubicin 150 | - |
| **Shayegi 2013** | SAL | Double induction | 2.800 | Daunorubicin 360 | - | - | - | - |
| **Terwjin 2012** | HOVON 42 | - | 1.400 | Idarubicin 36 | - | 12.000 | - | Amsacrine 360 |
| | HOVON 42a | - | 10.000 | Idarubicin 36 | - | 16.000 | - | Amsacrine 360 |
| | HOVON 29 | - | 1.400 | Idarubicin 36 | - | 12.000 | - | Amsacrine 360 |
| **Terwjin 2013** | HOVON 42a | - | 10.000 | Idarubicin 36 | - | 16.000 | - | Amsacrine 360 |
| **Willekens 2016** | CBF 2006 | 3+7 | 1.400 | Daunorubicin 180 | - | 18.000 | - | - |
| | CBF 2006 | - | 7.500 | Daunorubicin 285 | - | 18.000 | - | - |
| **Yin 2012** | MRC AML-15 | DA | 2.000 | Daunorubicin 150 | Gemtuzumab 3 | - | - | - |
| | MRC AML-15 | ADE | 2.000 | Daunorubicin 150 | Etoposide 500<br>Gemtuzumab 3 | - | - | - |
| | MRC AML-15 | FLAG-Ida | 10.000 | Idarubicin 24 | Gemtuzumab 3<br>Fludarabine 150 | - | - | - |
| **Yoon 2014** | - | 3+7 | 1.400 | Idarubicin 36 | - | 12.000 | - | - |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Zeijlemaker 2015** | HOVON | 3+7 | 1.400 | Idarubicin 36 | - | 12.000 | - | - |
| **Zhang 2013** | - | 3+7 | 700 | Idarubicin 36 | - | 12.000 | - | - |
| | - | IDAC | 4.400 | Idarubicin 36 | - | 12.000 | - | - |
| **Wei 2021** | - | 3+7 | 700 | Daunorubicin 180 | - | 9.000 | - | - |
| | - | IDAC | 6.440 | Daunorubicin 180 | - | 18.000 | - | - |

**Table S3**. Characteristics of patients according to treatment group after induction cycle.

| | **Overall n = 194** | **SDAC n = 149 (66.5%)** | **HDAC n = 45 (63.9%)** | ***P* value** | |
|---|---|---|---|---|---|
| **Age,** median (range) | 55 (16-73) | 57 (16-73) | 47 (19-61) | **<0.0001** | |
| **WBC,** $x10^9$/L, median (range) | 0.6 (14.8-435) | 17.8 (0.6-435.0) | 11.9 (1.1-289) | 0.70 | |
| **Hb,** g/dL, median (range) | 9.2 (3.9-14.9) | 9.1 (3.9-14.9) | 9.2 (4.2-12.8) | 0.25 | |
| **Plt,** $x10^9$/L, median (range) | 53 (1-281) | 53 (10-281) | 44 (1-216) | 0.36 | |
| **Peripheral blasts,** %, median (range) | 53 (0-100) | 52 (0-100) | 53 (0.98) | 0.86 | |
| **Secondary AML,** n (%) | 12 (6.2%) | 12 (8.0) | 0 (-) | | |
| **Karyotype,** n (%) | | | | | |
| *Favorable* | 33 (17.0) | 23 (15.4) | 10 (22.2) | 0.36 | 0.29 |
| *Normal* | 110 (56.7) | 87 (58.4) | 23 (51.1) | 0.40 | |
| *Intermediate, non-normal* | 26 (13.4) | 18 (12.1) | 8 (17.8) | 0.33 | |
| *Adverse* | 15 (7.7) | 11 (7.4) | 4 (8.9) | 0.75 | |
| *Lack of growth* | 10 (5.2) | 10 (6.7) | 0 (-) | 0.12 | |
| **Molecular genetics,** n (%) | | | | | |
| *NPM1-mutated* | 80 (41.2) | 66 (44.3) | 14 (31.1) | 0.12 | |
| *FLT3-ITD* | 41 (21.1) | 36 (24.2) | 5 (11.1) | 0.06 | |
| *CEBPA-bZIP* | 10 (5.2) | 7 (4.7) | 3 (6.7) | 0.70 | |
| **ELN 2010 risk groups,** n (%) | | | | | |
| *Favorable* | 90 (46.4) | 69 (46.3) | 21 (46.7) | 1.0 | 0.76 |
| *Intermediate-1* | 80 (41.2) | 63 (42.3) | 17 (37.8) | 0.61 | |
| *Intermediate-2* | 8 (4.1) | 5 (3.4) | 3 (6.7) | 0.39 | |
| *Adverse* | 16 (8.2) | 12 (8.1) | 4 (8.9) | 0.77 | |
| **ELN 2017 risk group,** n (%) | | | | | |
| *Favorable* | 102 (52.5) | 76 (51.0) | 26 (57.8) | 0.49 | 0.74 |
| *Intermediate* | 62 (32.0) | 49 (32.9) | 13 (28.9) | 0.72 | |
| *Adverse* | 24 (12.4) | 18 (12.1) | 6 (13.3) | 0.80 | |
| *Not assessable* | 6 (3.1) | 6 (4.0) | 0 (-) | 0.33 | |