Quantification of measurable residual disease using duplex sequencing in adults with acute myeloid leukemia

Laura W. Dillon,[1*] Jake Higgins,[2*] Hassan Nasif,[3] Megan Othus,[3] Lan Beppu,[4] Thomas H. Smith,[2] Elizabeth Schmidt,[2] Charles C. Valentine III,[2] Jesse J. Salk,[2] Brent L. Wood,[5] Harry P. Erba,[6] Jerald P. Radich[4,7#] and Christopher S. Hourigan[1,8,#]

[1]Laboratory of Myeloid Malignancies, Hematology Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD; [2]TwinStrand Biosciences, Seattle, WA; [3]Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA; [4]Clinical Research Division, Fred Hutchinson Cancer Center, Seattle, WA; [5]Department of Pathology and Laboratory Medicine, Children's Hospital Los Angeles, Los Angeles, CA; [6]Duke University School of Medicine, Durham, NC; [7]Division of Medical Oncology, Department of Medicine, University of Washington, Seattle, WA and [8]Myeloid Malignancies Program, National Institutes of Health, Bethesda, MD, USA

LWD and JH contributed equally as first authors.

#JPR and CSH contributed equally as senior authors.*

Correspondence: C.S. Hourigan
hourigan@nih.gov

Received: May 15, 2023.
Accepted: July 28, 2023.
Early view: August 3, 2023.

https://doi.org/10.3324/haematol.2023.283520

Abstract

The presence of measurable residual disease (MRD) is strongly associated with treatment outcomes in acute myeloid leukemia (AML). Despite the correlation with clinical outcomes, MRD assessment has yet to be standardized or routinely incorporated into clinical trials and discrepancies have been observed between different techniques for MRD assessment. In 62 patients with AML, aged 18-60 years, in first complete remission after intensive induction therapy on the randomized phase III SWOG-S0106 clinical trial (*clinicaltrials gov. Identifier: NCT00085709*), MRD detection by centralized, high-quality multiparametric flow cytometry was compared with a 29-gene panel utilizing duplex sequencing (DS), an ultrasensitive next-generation sequencing method that generates double-stranded consensus sequences to reduce false positive errors. MRD as defined by DS was observed in 22 (35%) patients and was strongly associated with higher rates of relapse (68% *vs.* 13%; hazard ratio [HR] =8.8; 95% confidence interval [CI]: 3.2-24.5; *P*<0.001) and decreased survival (32% *vs.* 82%; HR=5.6; 95% CI: 2.3-13.8; *P*<0.001) at 5 years. DS MRD strongly outperformed multiparametric flow cytometry MRD, which was observed in ten (16%) patients and marginally associated with higher rates of relapse (50% *vs.* 30%; HR=2.4; 95% CI: 0.9-6.7; *P*=0.087) and decreased survival (40% *vs.* 68%; HR=2.5; 95% CI: 1.0-6.3; *P*=0.059) at 5 years. Furthermore, the prognostic significance of DS MRD status at the time of remission for subsequent relapse was similar on both randomized arms of the trial. These findings suggest that next-generation sequencing-based AML MRD testing is a powerful tool that could be developed for use in patient management and for early anti-leukemic treatment assessment in clinical trials.

Introduction

Acute myeloid leukemia (AML) is a rare blood cancer diagnosed in approximately 20,000 Americans annually. While most patients treated with chemotherapy will achieve an initial complete remission (CR), less than one-third are expected to survive after 5 years.[1,2]

Measurable residual disease (MRD) is the presence of leukemia below the threshold set for remission by traditional clinical criteria but detectable with higher sensitivity approaches.[3] The presence of MRD is strongly associated with treatment outcomes.[4,5] However, despite being well established as correlated with the antileukemic effect of treatment interventions,[6-11] clinical implementation has been limited. While no standard technique is currently used for AML MRD testing,[12] multiple methodologies exist including detection of aberrant cell surface protein expression by multiparametric flow cytometry (MFC) or detection of genetic alterations by molecular assays.[13]

MFC has been widely used for AML MRD detection, but there are concerns that inter-laboratory variability and a lack of standardization could limit applicability of the technique on a broader scale.[14,15] MFC and next-generation sequencing (NGS) have been found to provide discordant

MRD results,[16,17] potentially capturing different residual cell populations. Furthermore, while MRD detection of certain highly prevalent genetic variants, including *FLT3* internal tandem duplications (*FLT3*-ITD) and *NPM1* insertions, by NGS has been shown to be strongly correlated with adverse clinical outcomes,[10,11,18] decentralized flow cytometry on the same patients was not prognostic.[11] There remains a need to compare AML MRD assessment using both centralized, high-quality MFC and ultra-sensitive NGS for detection of a broad range of variants in the same patients.

The SWOG Cancer Research Network S0106 study was an open-label randomized phase III clinical trial of adults aged 18-60 years with previously untreated *de novo* non-acute promyelocytic leukemia (non-APL) AML comparing standard induction therapy with daunorubicin (60 mg/m^2 intravenously [IV] day [d]1,2,3) and cytarabine (100 mg/m^2/d continuous infusion d1-7) ("DA") against the combination of daunorubicin (45 mg/m^2 d1,2,3), cytarabine (100 mg/m^2/d continuous infusion d1-7), and gemtuzumab ozogamicin (6 mg/m^2 d4) ("DA+GO"). Rates of cytomorphological CR (69% and 70%), 5-year relapse-free survival (RFS, 43% and 42%), and 5-year overall survival (OS, 46% and 50%) have previously been reported as not different between DA and DA+GO arms respectively.[19]

Utilizing samples and clinical data from patients treated on the S0106 trial, we explored the utility of MRD to predict treatment outcomes by both MFC and NGS. Bone marrow (BM) specimens obtained prior to treatment and at time of CR underwent centralized, prospective assessment of MRD using a three-tube, ten-color MFC assay.[20] Banked samples from a total of 67 patients were available at diagnosis and CR after first induction, and 62 patients with trackable variants identified using a 29-gene NGS panel at diagnosis underwent retrospective genomic analysis with error-corrected duplex sequencing (DS) for MRD at time of CR.

Methods

Patients

Archival BM aspirates or peripheral blood (PB) from 67 patients enrolled on the SWOG trial S0106 (*clinicaltrials gov. Identifier: NCT00085709*) were available for this study. A total of 62 patients were selected for MRD analysis if they (i) achieved a first morphological CR with protocol induction therapy, (ii) had both diagnosis and remission samples after first induction, (iii) had central flow cytometry results on their remission BM, and (iv) had a variant detected at diagnosis for tracking in remission. Samples described in this manuscript were collected at time of first morphologic CR, but if CR samples were not available the first sample collected after achieving CR was used. BM (n=56) and PB (n=6) remission samples were collected a median of 34 days (range, 25-162) post-randomization and a median of 0 days (range, -6 to 121) from clinically defined remission. The Institutional Review Board of the Fred Hutchinson Cancer Center gave ethical approval for this work, and patients were treated according to the Declaration of Helsinki.

Duplex sequencing

Retrospective targeted DNA sequencing of 29 genes recurrently mutated in adult AML was performed on genomic DNA (gDNA) collected from paired diagnostic and remission BM or PB samples utilizing the TwinStrand Duplex Sequencing™ AML-29 Panel (*Online Supplementary Table S1*). Non-error corrected sequencing was performed on diagnostic samples (500 ng gDNA) and error-corrected DS was performed on remission samples (1μg gDNA). DS was performed essentially as described[21] and further detailed in the *Online Supplementary Appendix*.

Bioinformatics

Alignment, duplex consensus sequence generation, and variant calling were performed as described.[21] For each patient, potential germline variants were identified and excluded from the analysis if the variant allele fraction (VAF) was ≥35% at both diagnosis and remission, or ≥40% at either time point and a gnomAD allele frequency ≥0.05. Somatic variants present at diagnosis were classified as potentially deleterious if computationally predicted as such and with a VAF ≥ 5% (≥1% for *FLT3*-ITD/*NPM1* insertions). Somatic variants in remission followed the same classification rules for deleterious impact and required an alternative depth of ≥2 (≥1 for *FLT3*-ITD/*NPM1* insertions detected at diagnosis). All remaining variants were manually curated for pathogenicity. MRD by NGS was defined using conditions previously identified as prognostic.[7,11]

Multiparametric flow cytometry

BM samples collected at diagnosis and remission were analyzed for MRD using a three-tube, ten-color MFC assay with a sensitivity of 0.1% in most cases; data and details of which were reported previously.[20]

Statistics

Morphologic complete remission was defined per contemporary consensus criteria definitions and required count recovery with absolute neutrophil count >1,000 and platelets >100,000. Time-to-event outcomes analyzed were OS (event=death), RFS (event=relapse or death) and time to relapse (TTR; event=relapse, death in remission a competing event). All outcomes were measured from date of morphologic remission to date of event, with patients without event censored at date of last contact. Associations between residual disease and outcomes were assessed using Cox regression models (cause-specific model for TTR); model discrimination was assessed using C-statistics.

Results

Patient characteristics

The median age of the 62 patients in this study was 48 years (range, 18-60) (Table 1). Thirty-two patients were randomized to DA and 30 patients to DA+GO. At 5 years, the rate of non-relapse mortality (NRM) was 10%, relapse was 33%, RFS was 57%, and OS was 64% for the entire cohort (Figure 1). Overall patient demographics and clinical outcomes of the 62 patients analyzed in this study align with the full S0106 clinical trial cohort (*Online Supplementary Table S2*).

Targeted sequencing analysis of diagnostic samples at an average raw sequencing read depth of 279x utilizing a 29 gene panel identified a total of 172 potentially deleterious variants across the 62 patients. Variants had a median VAF of 34% (range, 1.4-91.5) and were detected in 23 genes, with *FLT3* being the most frequently mutated (*Online Supplementary Table S3*; *Online Supplementary Figure S1*). Patients had a median of two variants detected at diagnosis (range 1-9) that could be tracked in remission.

Table 1. Patient clinical characteristics.

Covariate	Patient cohort
Patients, N	62
Randomized arm, N (%) DA DA+GO	 32 (52) 30 (48)
Age in years, median (range)	48 (18-60)
Sex, N (%) Female Male	 28 (45) 34 (55)
Performance status, N (%) 0-1 2-3	 58 (84) 11 (16
Cytogenetic risk, N (%) Favorable Intermediate Adverse Missing	 13 (23) 30 (54) 13 (23) 6
WBC x10^3/uL, median (range)	18.0 (0.2-214)
Platelets x10^3/uL, median (range)	48.5 (10-449)
Hemoglobin g/dL, median (range)	9.4 (3.5-13.6)
Race, N (%) Asian Black Native American/Alaskan Pacific Islander White Unknown	 1 (2) 4 (6) 1 (2) 0 54 (87) 2 (3)
Specimen for sequencing, N (%) Bone marrow Peripheral blood	 56 (90) 6 (10)

WBC: white blood cell count; DA: daunorubicin and cytarabine; GO: gemtuzumab ozogamicin.

Technical performance of duplex sequencing

Technical performance of the 29-gene DS assay was assessed on contrived mutation mixes *versus* healthy donor DNA. A single nucleotide variant mix containing 15 variants, an insertion-deletion mix containing four variants, and four separate serial dilutions of a *FLT3*-ITD/*NPM1* mutant mix were analyzed, with predicted VAF ranging from $1.0x10^{-2}$ to $3.9x10^{-6}$. Data combined from four replicate libraries per mix generated 135,065-142,707x mean duplex consensus molecular depth (from the 1.5 μg DNA input libraries), with max depths 186,645-196,896x. All expected variants were detected in the mutation mixes and the observed VAF were significantly correlated with the predicted VAF (r^2>0.99; *Online Supplementary Figure S2*). When the 21 spike-in mutation positions were assessed in the pure healthy donor DNA, a total of four mutant allele counts were detected out of a total duplex molecular depth of 2,993,429x at the 21 spike-in sites, for a combined mutation frequency of $1.3x10^{-6}$. The highest single background VAF at a spike-in site in the pure healthy donor DNA was $1.3x10^{-5}$.

Detection of residual variants in remission

DS of remission samples utilizing the same 29-gene panel at a median error-corrected duplex molecular depth of 27,996x (range, 11,958x-35,131x) identified 82 diagnostic variants remaining in remission, with a median VAF of 0.059% (range, 0.005-41.8) (*Online Supplementary Table S3*). Variants were detected in 18 genes, with *DNMT3A* being the most frequently mutated, followed by *NPM1* and *FLT3* (*Online Supplementary Figure S3*). Forty-three patients (69%) had at least one diagnostic variant detectable in remission, with a median of two residual variants per positive patient (range, 1-5). Residual diagnostic variants in remission had a median 2.60 (range, 0.06-3.96) log^{10} reduction in VAF. Not surprisingly, mutations in *DNMT3A* and *TET2*, genes commonly associated with clonal hematopoiesis, showed the least change in VAF between diagnosis and remission: median 1.23 (range, 0.06-3.31) and median 1.32 (range 1.23-2.29) log^{10} reduction, respectively. Mutations in *FLT3* showed the greatest change in VAF, median 3.12 (range, 1.5-3.8) log^{10} reduction.

Measurable residual disease as defined by flow cytometry

MFC analysis of BM collected at the time of remission using a three-tube, ten-color assay identified MRD in ten (16%) patients (Figure 2) and the median MRD level was 0.25% (range, 0.002-6.2%) (*Online Supplementary Table S4*). Patients who were MFC MRD-positive had increased rates of relapse (50% *vs.* 30% at year 5; hazard ratio [HR] =2.4; 95% confidence interval [CI]: 0.9-6.7; *P*=0.087) and decreased rates of RFS (40% *vs.* 61% at year 5; HR=2.2; 95% CI: 0.9-5.4; *P*=0.095) and OS (40% *vs.* 68% at year 5;

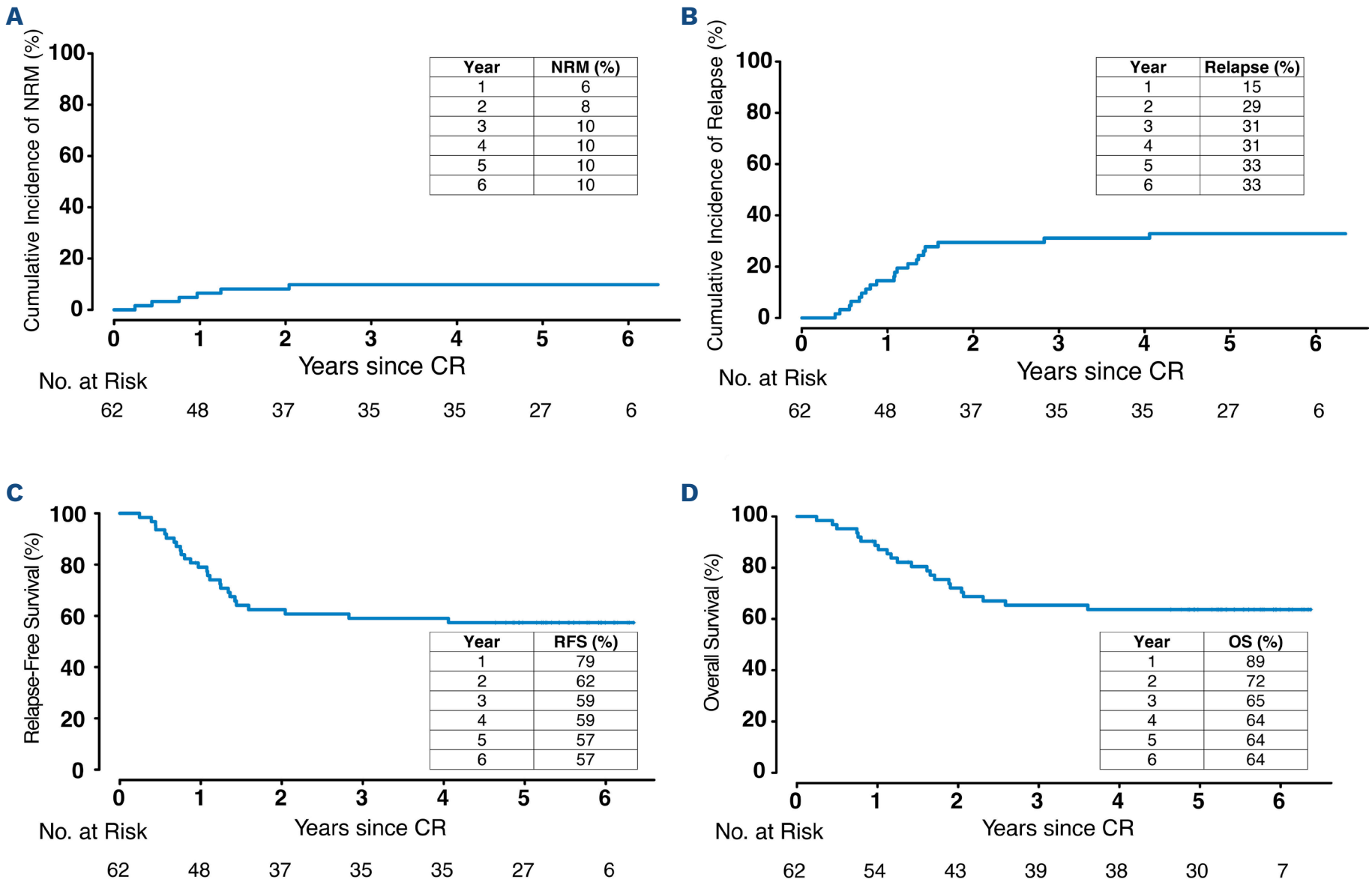


Figure 1. Clinical outcomes of S0106 acute myeloid leukemia patients analyzed for measurable residual disease. Rates of (A) non-relapse related mortality (NRM), (B) relapse, (C) relapse-free survival (RFS), and (D) overall survival (OS) are shown for the 62-patient cohort from the S0106 clinical trial analyzed for measurable residual disease (MRD) by duplex sequencing and multiparametric flow cytometry. CR: complete remission; No.: number.

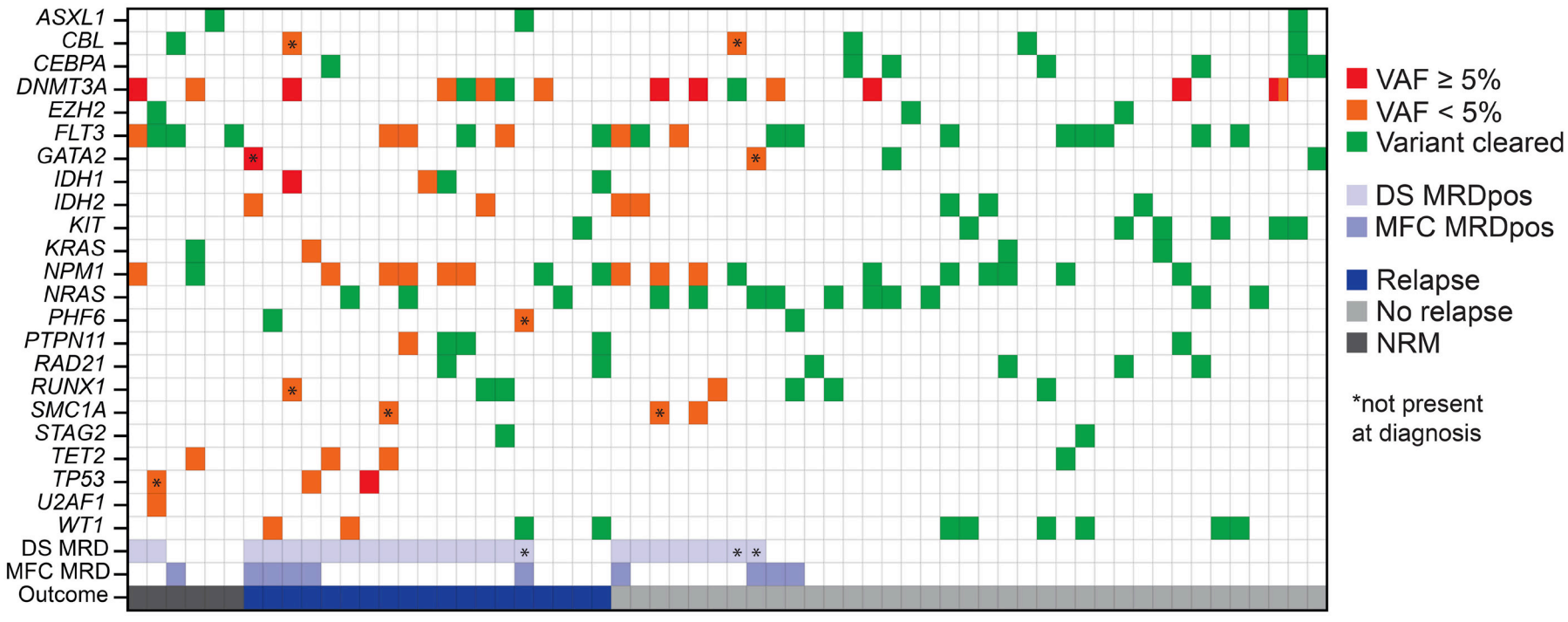


Figure 2. Mutational spectrum, measurable residual disease status, and clinical outcomes of patients in complete remission. The heatmap displays variants detected at diagnosis and the presence (divided into variant allele fraction [VAF] ≥ or <5%) or absence at the time of complete remission (CR) by duplex sequencing (DS), DS measurable residual disease (MRD) status, multiparametric flow cytometry (MFC) MRD status, and clinical outcome at 5 years (relapse, no relapse, or non-relapse mortality [NRM]). The presence of a mutation within a gene is denoted in the heatmap, with the color corresponding to the highest VAF within each gene per patient. Variants identified in remission that were not identified at diagnosis are also marked (*). pos: positive.

HR=2.5; 95% CI: 1.0-6.3; *P*=0.059) compared to patients that were MFC MRD-negative (Figure 3A; Table 2). While not statistically significant in this subset of S0106 patients, these results, including the magnitude of the HR, are in line with the significant findings previously published for the larger cohort of S0106 patients analyzed by MFC.[20]

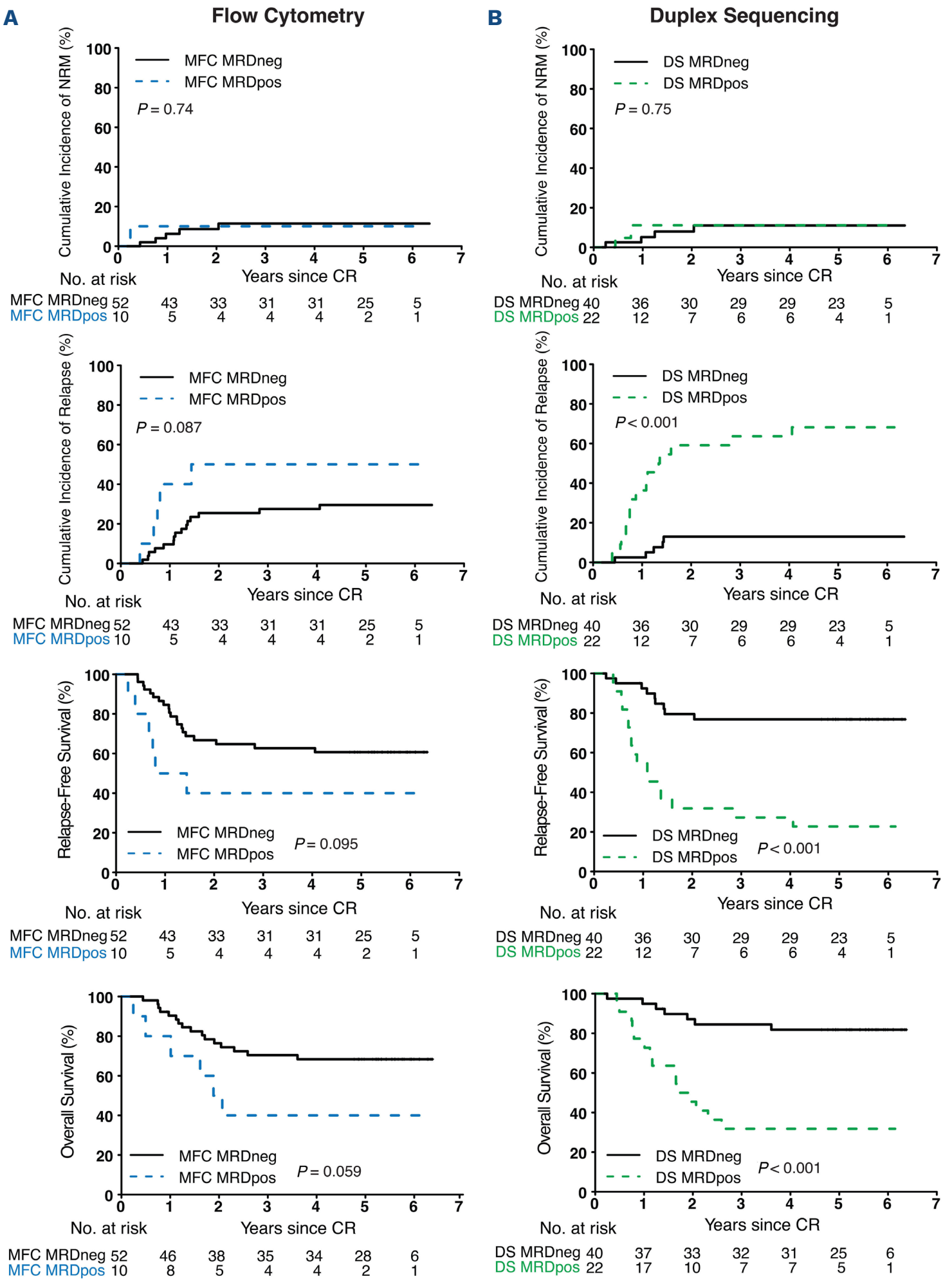


Figure 3. Impact of measurable residual disease status on clinical outcomes. Rates of non-relapse mortality (NRM), relapse, relapse-free survival, and overall survival are shown based on measurable residual disease (MRD) status as determine by (A) multiparametric flow cytometry (MFC) and (B) duplex sequencing (DS). pos: positive; neg: negative; No.: number; CR: complete remission.

Measurable residual disease as defined by detection of residual diagnostic variants by duplex sequencing

We defined DS test positivity utilizing criteria previously demonstrated to be prognostic for AML MRD by NGS,[7,11] which included non-*DTA* (*DNMT3A*, *TET2*, *ASXL1*) time-of-diagnosis mutations with a VAF ≥0.1% and/or an *FLT3*-ITD/*NPM1* VAF ≥0.01% (Figure 2). Using this definition, 22 patients (35%) were DS MRD-positive. Compared to MFC, DS MRD provided a superior prediction of clinical outcomes, such that patients who were DS MRD-positive had significantly increased rates of relapse (68% *vs.* 13% at year 5; HR=8.8; 95% CI: 3.2-24.5; *P*<0.001) and decreased rates of RFS (23% *vs.* 77% at year 5; HR=5.4; 95% CI: 2.4-12.3; *P*<0.001) and OS (32% *vs.* 82% at year 5; HR=5.6; 95% CI: 2.3-13.8; *P*<0.001) compared to patients that were DS MRD-negative (Figure 3B; Table 2).

Additional criteria for defining MRD by DS were also explored, including investigating the presence of any residual diagnostic variant and filtering based on VAF, gene, and VAF $\log^{10}$ reduction relative to diagnosis (*Online Supplementary Figure S4*; *Online Supplementary Table S5*). While a naïve definition of AML MRD as the detection of any residual diagnostic variant in remission was not associated with statistically significant difference in rates of relapse or survival (*Online Supplementary Figure S4*; *Online Supplementary Table S5*), the addition of VAF cutoffs, removal of mutations in genes associated with clonal hematopoiesis (*DTA*), and limiting calls to variants with no more than a $\log^{10}$ reduction of 2 between diagnosis and remission all resulted in statistically significant increased rates of relapse and decreased OS and RFS compared to patients testing negative, but none outperformed the criteria previously established as prognostic.

Measurable residual disease as defined by *de novo* detection of deleterious variants by duplex sequencing

We also explored the value of detecting AML-associated variants in remission that were not detected at the time of diagnosis. Utilizing the same variant filtering as defined above but agnostic to variant status at diagnosis, we identified 12 additional variants across nine patients with a median VAF of 0.24% (range, 0.08-15.1) (*Online Supplementary Table S3*). This resulted in three additional patients being defined as DS MRD-positive, for a total of 25 (40%) patients (Figure 2). Use of this NGS MRD definition agnostic to diagnostic variants provided a similar prediction of clinical outcomes to that of the initial prognostic criteria, such that patients who were DS MRD-positive had significantly increased rates of relapse (64% *vs.* 11% at year 5; HR= 8.7; 95% CI: 2.9-26.1; *P*<0.001) and decreased rates of RFS (28% *vs.* 78% at year 5; HR=4.8; 95% CI: 2.1-11.1; *P*<0.001) and OS (36% *vs.* 83% at year 5; HR=5.4; 95% CI: 2.1-13.8; *P*<0.001) compared to patients that were DS MRD-negative (*Online Supplementary Figure S4F*; *Online Supplementary Table S5*).

Comparison of measurable residual disease detection by multiparametric flow cytometry *versus* duplex sequencing

Next, we examined the differences between MFC and DS MRD calls. Of the 62 patients analyzed, five (8%) were called positive and 35 (56%) were called negative for MRD by both MFC and DS (Figure 4A). Five of the ten (50%) patients called MRD-positive by MFC were called negative by DS and 17 of the 22 (77%) patients called MRD-positive by DS were called negative by MFC.

Comparing clinical outcomes of the discordant cases revealed that 59% of patients called MFC MRD-negative/DS MRD-positive relapsed, while only 20% of patients called MFC MRD-positive/DS MRD-negative relapsed (Figure 4B). While patients defined as MRD-positive by both MFC and DS had the highest rate of relapse (80% at year 5), there was no significant difference in rates of relapse between DS MRD-positive/MFC-positive and DS MRD-positive/MFC-negative (80% *vs.* 65% at year 5; cause-specific *P*=0.59) or DS MRD-negative/MFC-positive and DS MRD-negative/MFC-negative (20% *vs.* 12% at year 5; cause-specific *P*=0.57), indicating DS MRD was the main driver of outcomes prediction (Figure 4C; *Online Supplementary Figure S5A*).

Looking closer at the disease burden in the discordant cases that experienced relapse, we found that the median VAF of variants identified in the DS MRD-positive/MFC-positive patients was 25 times higher than those identified in the DS MRD-positive/MFC-negative patients (1% *vs.* 0.04%). Additionally, five of the ten (50%) DS MRD-positive/MFC-negative patients that experienced relapse had a residual variant in *NPM1* detected, compared to none in the DS MRD-positive/MFC-positive patients.

Furthermore, the addition of MFC to the DS MRD definition did not significantly improve outcome predictions. While patients who were MFC- and/or DS MRD-positive had sig-

Table 2. Univariate cox regression model for associations between measurable residual disease definitions and clinical outcomes.

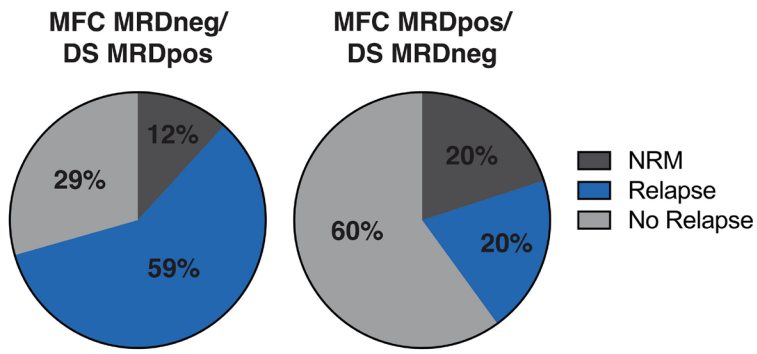
MRD definition	Relapse		Relapse-free survival		Overall survival	
	HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*
MFC	2.4 (0.9-6.7)	0.087	2.2 (0.9-5.4)	0.095	2.5 (1-6.3)	0.059
DS	8.8 (3.2-24.5)	<0.001	5.4 (2.4-12.3)	<0.001	5.6 (2.3-13.8)	<0.001
MFC+DS	7.8 (2.6-23.5)	<0.001	5.2 (2.2-12.5)	<0.001	6.2 (2.3-16.9)	<0.001

MFC: multiparametric flow cytometry; DS: duplex sequencing; HR: hazard ratio; CI: confidence interval

A

	MFC MRDneg (N=52)	MFC MRDpos (N=10)	Total
DS MRDneg	35 (67%)	5 (50%)	40 (64.5%)
DS MRDpos	17 (33%)	5 (50%)	22 (35.5%)

B



C

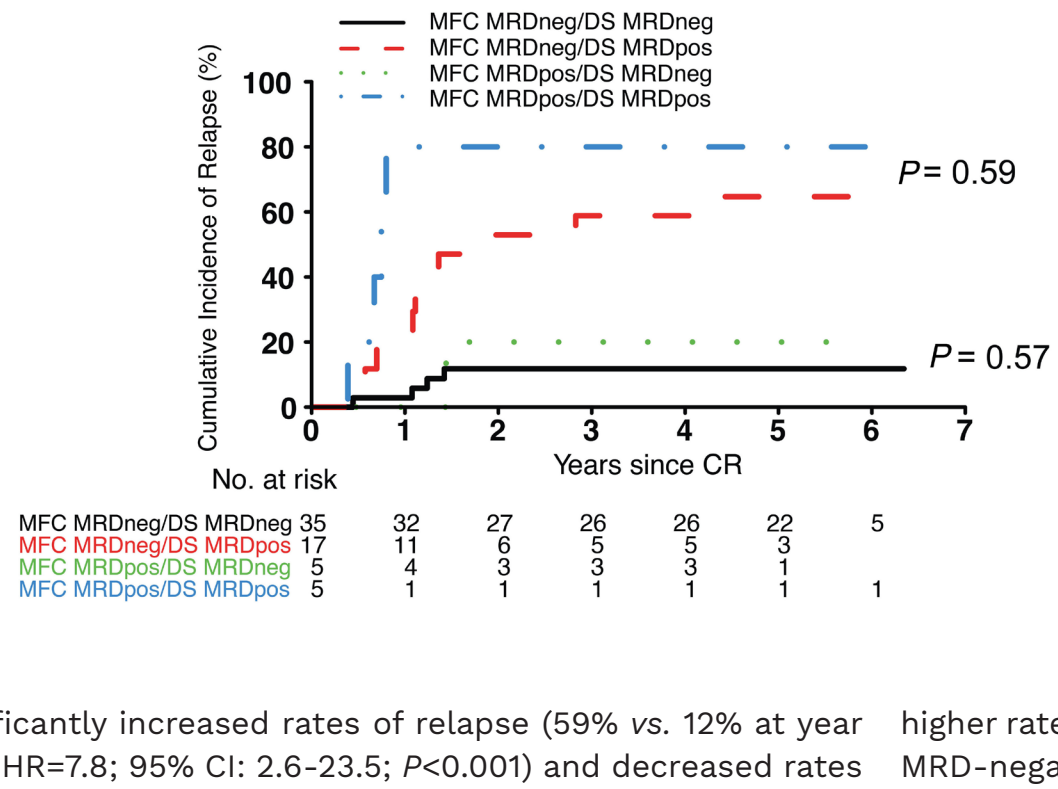


Figure 4. Analysis of discordant measurable residual disease results by duplex sequencing and flow cytometry. (A) Number and percentage of patients called measurable residual disease (MRD)-positive (pos) *versus* MRD-negative (neg) by duplex sequencing (DS) *versus* multiparametric flow cytometry (MFC). (B) Clinical outcomes (non-relapse mortality [NRM], relapse, or no relapse) of MFC MRD *versus* DS MRD discordant cases. (C) Rates of relapse for patients grouped by MRD status as defined by MFC MRD and DS MRD. CR: complete remission; No.: number.

nificantly increased rates of relapse (59% *vs.* 12% at year 5; HR=7.8; 95% CI: 2.6-23.5; *P*<0.001) and decreased rates of RFS (30% *vs.* 79% at year 5; HR=5.2; 95% CI: 2.2-12.5; *P*<0.001) and OS (37% *vs.* 85% at year 5; HR=6.2; 95% CI: 2.3-16.9; *P*<0.001) compared to patient who were MFC- and DS MRD-negative, this did not significantly differ from DS MRD alone (Table 2; *Online Supplementary Figure S5B*).

We assessed the ability of covariates to predict relapse in individual patients. Baseline clinical characteristics (including age, performance status, and cytogenetics) yielded a C-statistic of 0.66. Inclusion of MFC MRD status did not improve the discrimination of the model with a C-statistic of 0.67, while inclusion of DS MRD status did improve the model discrimination yielding a C-statistic of 0.77.

Impact of duplex sequencing measurable residual disease status and treatment regimen on clinical outcomes

Finally, we examined the impact of DS MRD status and patient randomization to DA *versus* DA+GO on clinical outcomes. In concordance with results from the full S0106 cohort,[19] the subset of 62 patients in this study showed no difference in rates of relapse between patients treated with DA *versus* DA+GO (35% *vs.* 31% at year 5; *P*=0.62) (Figure 5A). Adding information on DS MRD status showed that patients who were DS MRD-positive had significantly higher rates of relapse compared to patients that were DS MRD-negative regardless of the treatment regimen (DA: 60% *vs.* 12% at year 5; *P*=0.017, DA+GO: 86% *vs.* 13% at year 5; *P*<0.001) (Figure 5B). No difference was seen in rates of relapse between patients treated with DA *versus* DA+GO based on DS MRD status (DS MRD-positive: 60% *vs.* 86% at year 5; *P*=0.2; DS MRD-negative, 12% *vs.* 13%; *P*=0.98).

Discussion

MRD has been well established as a method for quantifying the antileukemic effect of interventional therapies, but implementation in the clinic has thus far been limited for AML. MFC has been widely used for AML MRD detection, but concerns exist over inter-laboratory variability which could hinder widespread applicability of this technique. NGS for AML MRD detection could be more amenable to decentralized clinical testing and has been shown to outperform decentralized flow cytometry in the context of *FLT3*-ITD- and *NPM1*-mutated AML.[11] Utilizing a subset of 62 AML patients treated on the S0106 phase III randomized trial of DA *versus* DA+GO induction chemotherapy, we compared the performance of MRD detection by high quality, centralized MFC and ultra-sensitive DS across a broad 29 gene panel to predict clinical response in first

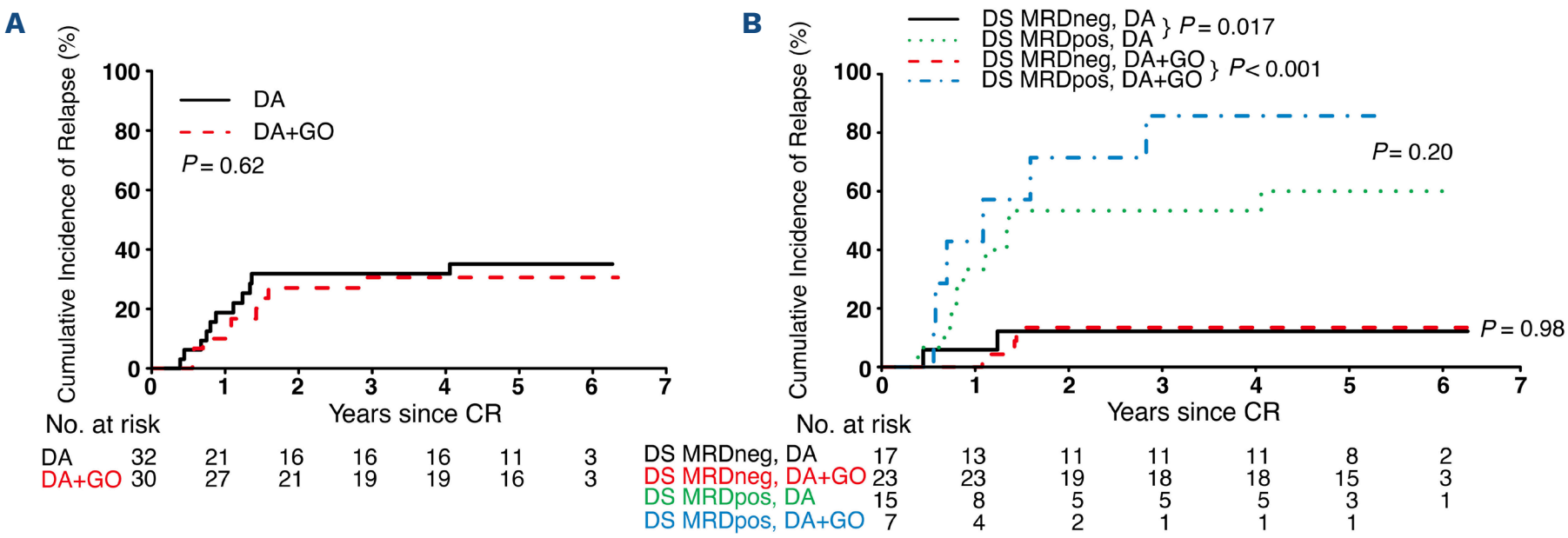


Figure 5. Impact of treatment randomization and duplex sequencing measurable residual disease status on relapse. Rates of relapse for patients as defined by (A) treatment randomization to daunorubicin and cytarabine (DA) *versus* daunorubicin, cytarabine, and gemtuzumab ozogamicin (DA+GO) and (B) treatment randomization (DA or DA+GO) and duplex sequencing (DS) measurable residual disease (MRD) status. pos.: positive; neg.: negative; No.: number; CR: complete remission.

remission and found the latter to broadly have superior outcome-predicting performance.

Application of NGS for AML MRD detection has varied across the literature, and questions remain regarding the impact of assay sensitivity, gene targets, variant status at diagnosis, and the applicability across patients with diverse baseline genetics.[22,23] In the 67 patients screened in this study, we found that 62 (93%) had at least one variant present at diagnosis in the gene panel examined that could be tracked by NGS. The mutations identified spanned 23 genes, representing a broad set of AML MRD targets. The highly sensitive DS assay detected residual mutations at some level in most patients, rendering the naïve designation of MRD positivity clinically uninformative. However, application of previously established, clinically relevant variant filtering conditions, including VAF thresholds well above the assay limit of detection and removal of less informative genes (*DNMT3A*, *TET2*, *ASXL1*) associated with clonal hematopoiesis,[24,25] was highly predictive of adverse clinical outcomes. These results highlight the importance of establishing informed guidelines for interpreting the presence of molecular MRD in the clinical setting. Additionally, we found that utilizing these filtering criteria remains highly predictive when agnostic to diagnostic variants. Therefore, the DS assay may have utility even when a diagnostic sample is not available. Future studies are needed to assess clinically relevant VAF thresholds at later treatment time points where residual disease may be present at a lower level.

In comparison to MFC, DS was significantly better at stratifying patients at risk of adverse clinical outcomes. Additional prognostic value was not seen when combining MRD detection by DS and MFC. Of the DS MRD-positive patients that relapsed, the median VAF of patients with MRD also detected by MFC was 25-times higher than MFC-negative patients and all relapses occurred within the first year, indicating that this subset of patients had a higher disease burden at the time of clinical remission. Of the DS MRD-positive patients that relapsed but were MFC-negative, 50% (n=5) had residual *NPM1* mutations, in contrast to none in the double-positive group. *NPM1*-mutated AML characteristically has absent/low CD34 expression with heterogeneity seen in the observed leukemia-associated immunophenotypes,[26,27] making it uniquely challenging to track by MFC. This combined with increased assay sensitivity could explain most of the discrepant results and improved prognostic power of DS.

The S0106 phase III clinical trial found that randomization of AML patients to DA *versus* DA+GO induction chemotherapy provided no significant difference in clinical outcomes. One potential value of MRD testing is to provide a surrogate endpoint to predict long-term patient response, allowing for faster drug development/approval and to identify patients in need of additional therapy versus those who do not. Recent trials have found that addition of GO to standard induction chemotherapy leads to deeper molecular responses and lower relapse rates in patients with *NPM1* mutated AML.[28,29] In this cohort we found that DS MRD was able to predict clinical relapse, with no significant difference for the prognostic implications of MRD status seen in patients who received DA *versus* DA+GO. Follow-up studies are needed to confirm the applicability of this technology as a definitive surrogate biomarker in clinical trial settings. Limitations of this study include (i) the small sample size, (ii) the retrospective nature of the DS MRD analysis, (iii) the comparison to an early generation MFC assay, and (iv) the age of the S0106 study potentially limiting comparability to contemporary AML standard of care. The findings of this study need to be confirmed in a larger cohort using prospective analysis by both DS and a more modern MFC MRD assay. While both flow cytometry and

molecular methods such as NGS can be used for AML MRD detection in both centralized and local settings, the level of expertise required for interpretation and the test performance characteristics differ.[30] Evidence-based recommendations support the use of molecular testing for a stable AML MRD target, in preference to flow cytometry, in situations where a validated test is available.[12,31] There are now multiple ongoing efforts to define appropriate targets, test requirements, interpretation, and clinical implications of AML MRD molecular testing.[31,32] Methods to suppress the false-positive error rates, such as DS, enable low-level variant discovery and could potentially expand the range of suitable targets for NGS-based AML MRD. In this study only five of the 67 (7%) patients screened were excluded due to lack of a mutation detected at diagnosis available for tracking by DS in the panel used. However, these patients did have cytogenetic abnormalities present. Whole exome or genome sequencing at diagnosis could inform individualized MRD panels, targeting a combination of recurrently mutated genes, novel variants, and structural alterations. Future work needs to be done exploring the use of patient personalized MRD targets to expand applicability to all patients.

In conclusion, we provide evidence that in a group of genetically diverse *de novo* adult AML patients randomized to DA *versus* DA+GO induction chemotherapy that ultra-sensitive detection of residual variants by DNA sequencing in the BM or PB at the time of first CR can outperform centralized, high-quality MFC in identifying patients at high risk of adverse clinical outcomes and predicting patient clinical response to treatment.

Disclosures

The National Heart, Lung, and Blood Institute receives research funding for the laboratory of CSH from the Foundation of the NIH AML MRD Biomarkers Consortium. MO consults for Merck and Biosight and serves on the data safety monitoring committee for Celgene, Glycomimetic, and Grifols. JH, THS, CCV, ES and JJS are employees and stockholders of TwinStrand Biosciences. BLW consults for Amgen and Kite Pharma. HPE discloses grants and other support from AbbVie, Agios Pharmaceuticals, ALX Oncology, Amgen, Daiichi Sankyo, FORMA Therapeutics, Forty Seven, Gilead Sciences, GlycoMimetics, ImmunoGen, Jazz Pharmaceuticals, MacroGenics, Novartis and PTC Therapeutics; has received research funding from AbbVie, Agios Pharmaceuticals, Bristol Myers Squibb, Celgene, Incyte Corporation, Jazz Pharmaceuticals and Novartis; is part of the speakers bureau of AbbVie; is on the independent review committee of AbbVie, Agios Pharmaceuticals, Astellas, Bristol Myers Squibb, Celgene, Daiichi Sankyo, Genentech, GlycoMimetics, Incyte Corporation, Jazz Pharmaceuticals and Kura Oncology. All other authors have no conflicts of interest to disclose.

Contributions

JPR conceived and designed the study. JH and ES performed and analyzed laboratory experiments. LB managed clinical samples. THS and CCV performed bioinformatic analysis. HN and MO performed statistical analysis. BLW performed flow cytometry analysis. HPE chaired the clinical trial. LWD performed genetic variant interpretation. LWD and CSH directed integrative analysis and wrote the original version of the manuscript. All authors contributed to reviewing and editing the manuscript and are accountable for the final version.

Funding

This work was supported in part by the Intramural Research Program of the National Heart, Lung, and Blood Institute; National Cancer Institute of the National Institutes of Health under award no. R44CA233381 (to JS); and National Cancer Institute CA175008, 180888, 180819, 233338, and 233381 (to JPR). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data-sharing statement

Raw FASTQ files are available in the NCBI Small Reads Archive (accession: PRJNA945188).

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022;72(1):7-33.
2. Dohner H, Wei AH, Appelbaum FR, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. Blood. 2022;140(12):1345-1377.
3. Hourigan CS, Karp JE. Minimal residual disease in acute myeloid leukaemia. Nat Rev Clin Oncol. 2013;10(8):460-471.
4. Short NJ, Fu C, Berry DA, et al. Association of hematologic response and assay sensitivity on the prognostic impact of measurable residual disease in acute myeloid leukemia: a systematic review and meta-analysis. Leukemia. 2022;36(12):2817-2826.
5. Short NJ, Zhou S, Fu C, et al. Association of measurable residual disease with survival outcomes in patients with acute myeloid leukemia: a systematic review and meta-analysis. JAMA Oncol. 2020;6(12):1890-1899.
6. Paras G, Morsink LM, Othus M, et al. Conditioning intensity and peritransplant flow cytometric MRD dynamics in adult AML. Blood. 2022;139(11):1694-1706.
7. 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. 2020;38(12):1273-1283.

8. Dillon R, Hills R, Freeman S, et al. Molecular MRD status and outcome after transplantation in NPM1-mutated AML. Blood. 2020;135(9):680-688.
9. Thol F, Gabdoulline R, Liebich A, et al. Measurable residual disease monitoring by NGS before allogeneic hematopoietic cell transplantation in AML. Blood. 2018;132(16):1703-1713.
10. Loo S, Dillon R, Ivey A, et al. Pretransplant FLT3-ITD MRD assessed by high-sensitivity PCR-NGS determines posttransplant clinical outcome. Blood. 2022;140(22):2407-2411.
11. Dillon LW, Gui G, Page KM, et al. DNA sequencing to detect residual disease in adults with acute myeloid leukemia prior to hematopoietic cell transplant. JAMA. 2023;329(9):745-755.
12. Heuser M, Freeman SD, Ossenkoppele GJ, et al. 2021 Update on MRD in acute myeloid leukemia: a consensus document from the European LeukemiaNet MRD Working Party. Blood. 2021;138(26):2753-2767.
13. Blachly JS, Walter RB, Hourigan CS. The present and future of measurable residual disease testing in acute myeloid leukemia. Haematologica. 2022;107(12):2810-2822.
14. Paiva B, Vidriales MB, 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.
15. 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. 2022;6(1):e676.
16. 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.
17. Patkar N, Kakirde C, Shaikh AF, et al. Clinical impact of panel-based error-corrected next generation sequencing versus flow cytometry to detect measurable residual disease (MRD) in acute myeloid leukemia (AML). Leukemia. 2021;35(5):1392-1404.
18. Grob T, Sanders MA, Vonk CM, et al. Prognostic value of FLT3-internal tandem duplication residual disease in acute myeloid leukemia. J Clin Oncol. 2023;41(4):756-765.
19. Petersdorf SH, Kopecky KJ, Slovak M, et al. A phase 3 study of gemtuzumab ozogamicin during induction and postconsolidation therapy in younger patients with acute myeloid leukemia. Blood. 2013;121(24):4854-4860.
20. Othus M, Wood BL, Stirewalt DL, et al. Effect of measurable ('minimal') residual disease (MRD) information on prediction of relapse and survival in adult acute myeloid leukemia. Leukemia. 2016;30(10):2080-2083.
21. Valentine CC, 3rd, Young RR, Fielden MR, et al. Direct quantification of in vivo mutagenesis and carcinogenesis using duplex sequencing. Proc Natl Acad Sci U S A. 2020;117(52):33414-33425.
22. Ghannam J, Dillon LW, Hourigan CS. Next-generation sequencing for measurable residual disease detection in acute myeloid leukaemia. Br J Haematol. 2020;188(1):77-85.
23. Walter RB, Ofran Y, Wierzbowska A, et al. Measurable residual disease as a biomarker in acute myeloid leukemia: theoretical and practical considerations. Leukemia. 2021;35(6):1529-1538.
24. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014;371(26):2488-2498.
25. Genovese G, Kahler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014;371(26):2477-2487.
26. Alcalay M, Tiacci E, Bergomas R, et al. Acute myeloid leukemia bearing cytoplasmic nucleophosmin (NPMc+ AML) shows a distinct gene expression profile characterized by up-regulation of genes involved in stem-cell maintenance. Blood. 2005;106(3):899-902.
27. Verhaak RG, Goudswaard CS, van Putten W, et al. Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. Blood. 2005;106(12):3747-3754.
28. Kapp-Schwoerer S, Weber D, Corbacioglu A, et al. Impact of gemtuzumab ozogamicin on MRD and relapse risk in patients with NPM1-mutated AML: results from the AMLSG 09-09 trial. Blood. 2020;136(26):3041-3050.
29. Lambert J, Lambert J, Nibourel O, et al. MRD assessed by WT1 and NPM1 transcript levels identifies distinct outcomes in AML patients and is influenced by gemtuzumab ozogamicin. Oncotarget. 2014;5(15):6280-6288.
30. Freeman SD, Hourigan CS. MRD evaluation of AML in clinical practice: are we there yet? Hematology Am Soc Hematol Educ Program. 2019;2019(1):557-569.
31. Hourigan CS. Achieving MRD negativity in AML: how important is this and how do we get there? Hematology Am Soc Hematol Educ Program. 2022;2022(1):9-14.
32. Wong ZC, Dillon LW, Hourigan CS. Measurable residual disease in patients undergoing allogeneic transplant for acute myeloid leukemia. Best Pract Res Clin Haematol. 2023;36(2):101468.