Measurable residual disease at induction redefines partial response in acute myeloid leukemia and stratifies outcomes in standard- risk patients without NPM1 mutations

Freeman, Sylvie; Hills, Robert; Virgo, Paul; Khan, Naeem; Couzens, S.; Dillon, Richard; Gilkes, Amanda; Upton, Laura; Nielsen, Ove Juul ; Cavenagh, James D; Jones, Gail; Khwaja, Asim; Cahalin, Paul; Thomas, Ian; Grimwade, David; Burnett, Alan K; Russell, Nigel H

DOI:
10.1200/JCO.2017.76.3425

Document Version
Publisher's PDF, also known as Version of record

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Measurable Residual Disease at Induction Redefines Partial Response in Acute Myeloid Leukemia and Stratifies Outcomes in Patients at Standard Risk Without *NPM1* Mutations

Sylvie D. Freeman, Robert K. Hills, Paul Virgo, Naeem Khan, Steve Couzens, Richard Dillon, Amanda Gilkes, Laura Upton, Ove Juul Nielsen, James D. Cavenagh, Gail Jones, Asim Khwaja, Paul Cahalin, Ian Thomas, David Grimwade,† Alan K. Burnett, and Nigel H. Russell

Author affiliations and support information (if applicable) appear at the end of this article.

†Deceased.

Published at jco.org on March 30, 2018.

Written on behalf of the UK National Cancer Research Institute AML Working Group.

S.D.F. and R.K. H. contributed equally to this work.

Corresponding author: Sylvie D. Freeman, MD, PhD, Birmingham Medical School, Vincent Drive, Edgbaston, Birmingham, United Kingdom; e-mail: s.freeman@bham.ac.uk.

0732-183X/18/3699-1/$20.00

ABSTRACT

Purpose

We investigated the effect on outcome of measurable or minimal residual disease (MRD) status after each induction course to evaluate the extent of its predictive value for acute myeloid leukemia (AML) risk groups, including *NPM1* wild-type (wt) standard risk, when incorporated with other induction response criteria.

Methods

As part of the NCRI AML17 trial, 2,450 younger adult patients with AML or high-risk myelodysplastic syndrome had prospective multiparameter flow cytometric MRD (MFC-MRD) assessment. After course 1 (C1), responses were categorized as resistant disease (RD), partial remission (PR), and complete remission (CR) or complete remission with absolute neutrophil count $<$ 1,000/μL or thrombocytopenia $<$ 100,000/μL (CRi) by clinicians, with CR/CRi subdivided by MFC-MRD assay into MRD+ and MRD−. Patients without high-risk factors, including *Flt3* internal tandem duplication wt/−*NPM1*-wt subgroup, received a second daunorubicin/cytosine arabinoside induction; course 2 (C2) was intensified for patients with high-risk factors.

Results

Survival outcomes from PR and MRD+ responses after C1 were similar, particularly for good- to standard-risk subgroups (5-year overall survival [OS], 27% RD *v* 46% PR *v* 51% MRD+ *v* 70% MRD−; $P < .001$). Adjusted analyses confirmed significant OS differences between C1 RD versus PR/MRD+ but not PR versus MRD+. CRi after C1 reduced OS in MRD+ (19% CRi *v* 45% CR; $P = .001$) patients, with a smaller effect after C2. The prognostic effect of C2 MFC-MRD status (relapse: hazard ratio [HR], 1.88 [95% CI, 1.50 to 2.36], $P < .001$; survival: HR, 1.77 [95% CI, 1.41 to 2.22], $P < .001$) remained significant when adjusting for C1 response. MRD positivity appeared less discriminatory in poor-risk patients by stratified analyses. For the *NPM1*-wt standard-risk subgroup, C2 MRD+ was significantly associated with poorer outcomes (OS, 33% *v* 63% MRD−, $P = .003$; relapse incidence, 89% when MRD+ $\geq$ 0.1%); transplant benefit was more apparent in patients with MRD+ (HR, 0.72; 95% CI, 0.31 to 1.69) than those with MRD− (HR, 1.68 [95% CI, 0.75 to 3.85]; $P = .16$ for interaction).

Conclusion

MFC-MRD can improve outcome stratification by extending the definition of partial response after first induction and may help predict *NPM1*-wt standard-risk patients with poor outcome who benefit from transplant in the first CR.

J Clin Oncol 36. © 2018 by American Society of Clinical Oncology

ASSOCIATED CONTENT

Appendix
DOI: https://doi.org/10.1200/JCO.2017.76.3425

Data Supplement
DOI: https://doi.org/10.1200/JCO.2017.76.3425

DOI: https://doi.org/10.1200/JCO.2017.76.3425

INTRODUCTION

In acute myeloid leukemia (AML), failure to achieve morphologic complete remission (CR) after a first cycle of induction in previously untreated patients is an established independent prognostic factor from earlier studies.[1-3] Thus, morphologic response at this time point is often incorporated with genetic and pretreatment clinical parameters to guide further therapy,[4] including second induction courses, choice of consolidation, and whether intensification from allogeneic stem-cell transplantation (SCT) may be appropriate in

otherwise intermediate-risk patients. Despite morphologic response criteria being standard, a different approach for measuring response has been proposed[5,6] owing to the independent prognostic value from measurable or minimal residual disease (MRD) assays when discrepant with morphology,[7-9] or in CR[10-12] and the equivalent poor outcomes between MRD positivity and active-disease pre-myeloablative SCT.[13,14]

Studies have shown the prognostic value of MRD monitoring by polymerase chain reaction (PCR) for patients with validated molecular targets, usually after two courses of chemotherapy.[11,12,15] Multiparametric flow cytometry MRD (MFC-MRD) may identify, as early as after course 1 (C1), patients with a poorer response despite achieving CR and is an assay that can be applied across AML genetic subgroups.[12,16-20] There are, however, insufficient data to ascertain the relative prognostic effect of MFC-MRD positivity in CR post-C1 compared with morphologic active disease; it is feasible that the outcomes of patients with detectable MRD resemble those of refractory patients who achieve the cytoreduction criteria for a morphologic partial remission (PR).[21,22] Evaluating this will help refine which response categories are the most useful prognostic surrogate end points to assess effectiveness of the first induction course.

It is also uncertain for patients who complete a second chemotherapy course whether the quality of response after C1, with inclusion of MFC-MRD assessment, adds prognostic information to CR-MRD status after course 2 (C2). The value of MFC-MRD status to differentiate outcome at either time point is likely to be heterogeneous between established risk subgroups due to disease, treatment, and assay factors, but the extent of this has not been established.

Treatment decisions, including predicting the benefit of SCT, are particularly challenging for the standard-risk subgroup. MFC-MRD assays are most likely to influence therapeutic choices for *NPM1*-wild type (wt) patients of standard risk, following data indicating post-induction reverse transcriptase, quantitative PCR (RT-qPCR) of blood-mutated transcripts reliably predicts outcome for patients with *NPM1* mutation.[23,24] Thus, there is a specific need to define the usefulness of MFC-MRD for risk stratification in this subgroup.

In this study, we aimed to determine the prognostic effect of MFC-MRD measurement incorporated into response assessment after induction courses for the different risk subgroups, including *NPM1*-wt patients at standard risk, in a large cohort of younger patients with AML who had undergone intensive treatment in the National Cancer Research Institute (NCRI) AML17 trial.

METHODS

Patients

Patients were enrolled in the NCRI AML17 trial (ISRCTN Registry No. 55675535) from April 6, 2009, to December 31, 2014. A list of treatments is provided in Appendix Fig. A1 (online only).

The AML17 protocol was designed primarily for younger patients, generally age $<$ 60 years. Patients with high-risk myelodysplastic syndrome, which was defined as $>$ 10% marrow blasts at diagnosis, and secondary AML were eligible. Patients with acute promyelocytic leukemia were not included in this MRD study. After first induction, patients were defined by risk of relapse, using a validated score comprising cytogenetics, WBC count, age, secondary disease, morphologic response to C1[25,26] and *FLT-3* internal tandem duplication(ITD)/*NPM1* mutation status.

Morphologic-based response criteria were as follows: (1) CR, $<$ 5% blasts in a cellular bone marrow with count recovery, CRi if 5% blasts but best response was with neutropenia $<$ 1,000/μL or thrombocytopenia $<$ 100,000/μL; (2) partial remission (PR), decrease of pretreatment bone marrow blast percentage by at least 50% to 5% to 15% in a cellular marrow (hematologic recovery not required)[1]; and (3) resistant disease (RD), $>$ 15% marrow blasts (patients surviving at least 7 days after completion of treatment). Responses were classified by centers.

Patients designated as favorable or at standard risk received the second daunorubicin/cytosine arabinoside course and were then randomized to receive either 1 or 2 courses of high-dose cytosine arabinoside. High-risk patients were offered a randomization between FLAG-Ida or daunorubicin/clofarabine with the intention of eventually proceeding to allogeneic stem-cell transplantation (SCT). *FLT3*-ITD mutant patients were directed to the lestaurtinib randomization until 2012.

The trial was sponsored by Cardiff University, approved by Wales-REC3 and conducted in accordance with the Declaration of Helsinki.

Multiparameter Flow Cytometry Detection of MRD

Samples for multiparameter flow cytometry (MFC)-MRD were requested at baseline (bone marrow and/or blood) and following each course (bone marrow). A summary of sample logistics and processing is provided in the Data Supplement. MFC-MRD analysis was performed centrally, using standardized gating strategy that screened for "different-from-normal" leukemia-associated-immunophenotypes (LAIPs) on blasts pretreatment and tracked these (approximately 0.02% to 0.05% sensitivity thresholds) but also applied the different-to-normal approach in follow-up samples to detect changes in blast LAIPs (approximately 0.05% to 0.1% sensitivity threshold). In this study, only samples for which there were pretreatment LAIPs to monitor could be reported as MFC-MRD negative, whereas samples with any level of MRD detected above a diagnostic LAIP or different-from-normal follow-up LAIP threshold were reported as MFC-MRD positive. Clinicians were not informed of MFC-MRD results.

Statistical Analysis

All end points were based on the revised criteria of the International Working Group for Diagnosis.[21] Survival percentages were calculated using the Kaplan-Meier method with cumulative incidence of relapse calculated using competing-risks methodology. Baseline characteristics were compared using χ^2 or Mantel-Haenszel tests, with continuous variables compared using the Wilcoxon rank-sum test. Time-to-event outcomes were compared using log-rank tests and Cox regression. Outcomes are reported as effect sizes with 95% confidence intervals; significance was set at $P < .05$. Stratified analyses used stratified log-rank tests and are displayed as forest plots with tests for interaction using standard methodology.[27] Comparison of transplantation versus not was analyzed using the method of Mantel and Byar to mitigate immortal time bias. Median follow-up for survival was 39.0 months (range, 1.0 to 80.5 months).

RESULTS

Induction Response by Morphology and MFC-MRD: Patient Characteristics

Between 2009 and 2014, 6,539 samples (bone marrow [BM] or peripheral blood at diagnosis, BM post-treatment courses) from 2,450 patients with non-acute promyelocytic leukemia recruited to AML17 were prospectively analyzed for MFC-MRD (Appendix Fig A2, online only). Among patients in CR post-C1, the presence of MRD data were associated with secondary AML, and the absence of an *NPM1* mutation (reflecting the prioritizing of BM for RT-qPCR monitoring of *NPM1* mutations[23] during the second phase of the trial); survival at 5 years was 52% (with MRD data) versus 50%

Table 1. Characteristics of Study Population by Response

Characteristic	Post-C1 Response (n = 1,443)							Post-C2 MRD status in CR (n = 806)		
	MRD−	MRD+	PR	RD	*P* (MRD− *v* MRD+)	*P* (four categories)	*P* (MRD+ *v* PR)	MRD−	MRD+	*P*
All patients	446 (30.9)	577 (40.0)	197 (13.7)	223 (15.5)				503 (62.4)	303 (37.6)	
Age, years										
16-29	48 (10.8)	67 (11.6)	22 (11.2)	26 (11.7)				67 (13.3)	29 (9.6)	
30-39	59 (13.2)	55 (9.5)	15 (7.6)	24 (10.8)				64 (12.7)	33 (10.9.)	
40-49	104 (23.3)	135 (23.4)	36 (18.3)	39 (17.5)				122 (24.3)	62 (20.5)	
50-59	145 (32.5)	187 (32.4)	75 (38.1)	68 (30.5)				168 (33.4)	99 (32.7)	
≥ 60	90 (20.2)	133 (23.1)	49 (24.9)	66 (29.6)				82 (16.3)	80 (26.4.)	
Age, years, median (range)	50 (16-71)	51 (16-72)	53 (16-72)	53 (17-73)	.3*	.02*	.3*	49 (16-72)	53 (16-72)	.002*
Sex					.06	.03†	1.0			.006
Female	220 (49.3)	250 (43.3.)	85 (43.1.)	91 (40.8.)				248 (49.3.)	119 (39.3.)	
Male	226 (50.7)	327 (56.7)	112 (56.9)	132 (59.2.)				255 (50.7.)	184 (60.7.)	
Diagnosis					.007	< .001	.5			0.12
De novo	410 (91.9.)	492 (85.3)	162 (82.2.)	193 (86.5.)				459 (91.3.)	265 (87.5.)	
Secondary	18 (4.0.)	52 (9.0.)	23 (11.7)	25 (11.2)				25 (5.0)	26 (8.6.)	
High-risk MDS	18 (4.0)	33 (5.7)	12 (6.1)	0 (0)				19 (3.7)	12 (4.0)	
WHO performance status					1.0†	.4†	.8†			0.7†
0	314 (70.4)	413 (71.6)	138 (70.1)	148 (66.4)				359 (71.4)	219 (72.3)	
1	115 (25.8)	138 (23.9)	53 (26.9)	65 (29.1)				124 (24.7)	72 (23.8)	
2	11 (2.5)	14 (2.4)	4 (2.0)	6 (2.7)				13 (25.8)	9 (3.0)	
3	5 (1.1)	12 (2.1)	2 (1.0)	4 (1.8)				6 (1.2)	3 (1.0)	
4	1 (0.2)	0 (0)	0 (0)	0 (0)				1 (0.1)	0 (0)	
WBC count (× 10^9/L)										
0-9.9	219 (49.1)	301 (52.2)	114 (57.9)	112 (50.2)				260 (51.7)	155 (51.2)	
10-49.9	156 (35.0)	186 (32.2)	47 (23.9)	55 (24.7)				166 (33.0)	88 (29.0)	
50-99.9	42 (9.4)	45 (7.8)	21 (10.7)	26 (11.7)				46 (9.1)	29 (9.6)	
≥ 100	29 (6.5)	45 (7.8)	15 (7.6)	30 (13.5)				31 (61.6)	31 (10.2)	
Median WBC (× 10^9/L) (range)	10.6 (0.3-319.6)	8.9 (0.4-456.0)	6.0 (0.4-430.0)	9.9 (0.6-430.0)	.3*	.3*	.13*	9.3 (0.7-275.2)	9.5 (0.4-430.0)	0.5*
Cytogenetics					< .001†	< .001†	< .001†			< .001
Favorable	82 (18.4)	75 (13.0)	8 (4.1)	3 (1.3)				94 (18.7)	18 (5.9)	
Intermediate	322 (72.2)	386 (66.9)	126 (64.0)	130 (58.3)				349 (69.4)	208 (41.4)	
Adverse	27 (6.1)	87 (15.1)	53 (26.9)	79 (35.4)				36 (7.2)	59 (11.7)	
Unknown	15 (3.3)	28 (4.9)	10 (5.1)	11 (4.9)				24 (4.8)	17 (3.4)	
FLT3-ITD/*NPM1c* status					.001	< .001	.05			.005
ITD WT, *NPM1c* WT	229 (51.3)	356 (61.7)	129 (65.5)	174 (78.0)				274 (54.5)	192 (63.4)	
ITD WT, *NPM1c* mutant	99 (22.2)	88 (15.3)	18 (9.1)	8 (3.6)				102 (20.3)	39 (12.9)	
ITD mutant, *NPM1c* WT	22 (5.0)	42 (7.3)	17 (8.6)	29 (13.0)				27 (5.4)	24 (7.9)	
ITD mutant, *NPM1c* mutant	63 (14.1)	63 (10.9)	12 (6.1)	5 (2.2)				66 (13.1)	28 (9.2)	
Unknown	33 (7.4)	28 (4.9)	21 (10.7)	7 (3.1)				34 (6.8)	20 (6.6)	
Post-C1 response	407 (91.3)	528 (91.5)			.9					< .001†; CR *v* CRi *P* = .8
CR	39 (8.7)	49 (8.5)						412 (81.9)	219 (72.3)	
CRi								37 (7.4)	18 (5.9)	
PR								26 (5.2)	33 (10.9)	
RD								22 (4.4)	28 (9.2)	
								6 (1.2)	5 (1.7)	
Post-C1 risk score					< .001†	< .001†	< .001†			< .001†
Good	86 (19.3)	85 (14.7)	11 (5.6)	3 (1.3)				102 (20.3)	30 (9.9)	
Standard	253 (56.7)	267 (46.3)	42 (21.3)	37 (16.6)				267 (53.1)	131 (43.2)	
Poor risk	102 (22.9)	218 (37.8)	138 (70.1)	167 (74.9)				132 (26.2)	141 (46.5)	
Not assessable	6 (1.3)	6 (1.0)	6 (3.0)	16 (7.2)				2 (0.4)	1 (0.3)	

(continued on following page)

Table 1. Characteristics of Study Population by Response (continued)

Characteristic	Post-C1 Response (n = 1,443)						Post-C2 MRD status in CR (n = 806)		
	MRD−	MRD+	PR	RD	*P* (MRD− *v* MRD+)	*P* (four categories)	*P* (MRD+ *v* PR)	MRD−	MRD+	*P*
No. of chemotherapy cycles (excludes patients at poor risk)					n/a	n/a	n/a			n/a
3	95 (21.3)	128 (22.2)	12 (6.1)	0 (0)				113 (22.5)	64 (21.1)	
4	117 (26.2)	116 (20.1)	10 (5.1)	1 (0.4)				139 (27.6	49 (16.2)	
Not randomized	139 (31.2)	130 (22.5)	37 (18.8)	56 (25.1)				128 (25.4)	56 (18.5)	
Allogeneic SCT					.07	< .001†	.01			.04
Any	142 (31.8)	215 (37.3)	102 (51.8)	89 (39.9)				193 (38.4)	139 (45.9)	
CR1	82 (18.4)	128 (22.2	75 (38.1	71 (31.8)				114 (22.7)	91 (30.0)	
CR2	38 (8.5)	68 (11.8)	4 (2.0)	2 (0.9)				54 (10.7)	32 (10.6)	

NOTE. Data given as No. (%) unless otherwise specified. χ^2 test unless otherwise specified.
Abbreviations: C1, course 1; C2, course 2; CR, complete remission; CR1, first complete remission; CRi, complete remission with absolute neutrophil count < 1,000/μL or thrombocytopenia < 100,000/μL; MDS, myelodysplastic syndrome; MRD, measurable residual disease; n/a, not assessed; PR, partial response (≥ 50% reduction in blast numbers with 5% to 15% residual blasts); RD, resistant disease (< 50% reduction in blast numbers with > 15% residual blasts); SCT, stem-cell transplant; WHO, World Health Organization.
*Wilcoxon rank sum test/Spearman correlation; not accessed as 3 versus 4 courses randomized.
†Mantel-Haenszel test for trend.

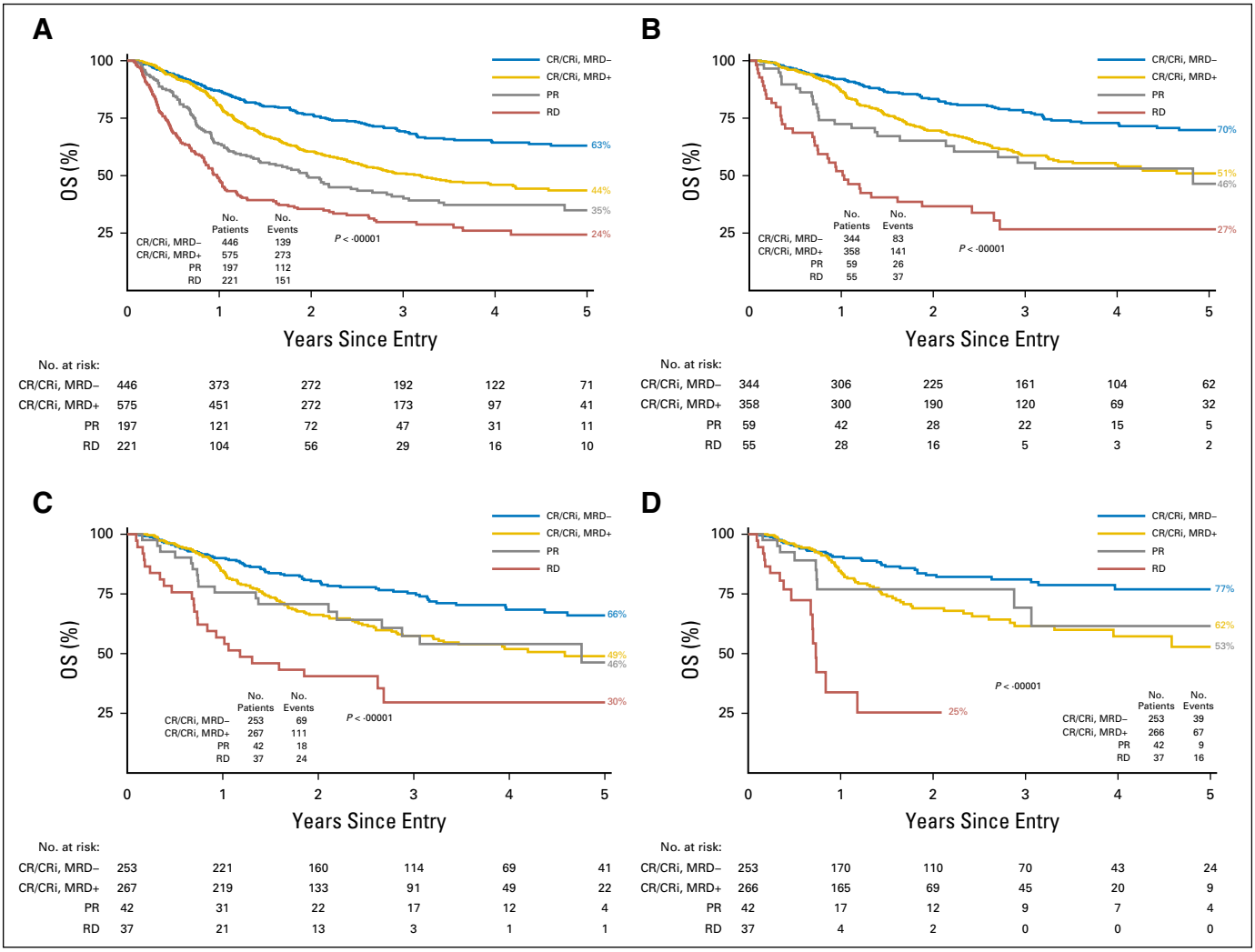

Fig 1. Overall survival (OS) according to response status after course 1. (A) All patients. (B) Patients at good and standard risk (patients known to be at poor risk excluded). (C) Patients at standard risk. (D) OS for patients at standard risk censored at allogeneic stem-cell transplantation. CR, complete remission; CRi, complete remission with absolute neutrophil count < 1,000/μL or thrombocytopenia < 100,000/μL; MRD, measurable residual disease; PR, partial remission; RD, resistant disease.

(without MRD data). In adjusted analyses, the presence of MRD data was not associated with survival (hazard ratio [HR], 0.99 [95% CI, 0.84 to 1.16]; $P = .9$).

Post-C1, 1,443 patients contributed data; 420 were refractory by morphology (n = 197 RD; n = 223 PR) and 1,023 (70.9%) achieved CR/CRi with MFC-MRD data (n = 446 MFC-MRD negative [MRD−]; n = 577 MFC-MRD positive [MRD+]). After C2, 806 patients were in CR/CRi with MFC-MRD data (n = 503 MRD−; n = 303 MRD+).

The clinical characteristics of patients according to response post-C1 and MRD status for patients in CR/CRi post-C1 or C2 are listed in Table 1. There was a significant association between responses post-C1 or C2 and cytogenetic group; however, count recovery post-C1 was not significantly associated with MRD after either course.

Outcome Comparison for Morphologic Response and MFC-MRD Status After C1

We evaluated overall survival (OS) by C1 response status. Five-year OS for all enrolled in AML17 excluding early deaths was 52% for those achieving CR/CRi versus 31% for refractory patients ($P < .001$). MRD status in CR/CRi versus PR or RD further differentiated 5-year survival outcomes (Fig 1A). A PR or MRD+ response gave intermediate survival at 5 years. Survival rates appeared equivalent between these two responses for the patients at good or standard risk; 5-year OS for MRD− versus MRD+ versus PR versus RD were 63% versus 44% versus 35% versus 24%, respectively, for all patients; 70% versus 51% versus 46% versus 27%, respectively, when patients at poor risk were excluded (Fig 1B); and 66% versus 49% versus 46% versus 30%, respectively, for standard risk alone ($P < .001$ for all analyses; Fig 1C). Similar results were observed for survival censored at SCT (Fig 1D; Fig A3A, online only) and also for *NPM1*-wt patients at standard risk (Fig A3B and A3C).

Adjusted analyses confirmed significant survival differences between RD and PR/MRD+ but not between PR and MRD+ for patients at good or standard risk (RD *v* PR/MRD+: OS HR, 2.28 [95% CI, 1.38 to 3.75]; $P < .001$; PR vs MRD+: HR, 1.32; $P = .4$) and for *NPM1*-wt patients at standard risk (RD *v* PR/MRD+: OS HR, 2.13 [95% CI, 1.21 to 3.75]; $P = .008$; PR vs MRD+: HR, 1.18, $P = .6$). Results were similar when censored at SCT (Table 2).

Thus, the prognostic effect from morphologic response criteria after first induction was restricted to RD in the good and standard-risk subgroups when MFC-MRD status was incorporated into response assessment.

Only 25 patients were refractory by morphology post-C1 but MRD− (n = 22 PR; n = 3 RD) with 61% 3-year and 49% 5-year OS. Seven of 577 MRD+ patients were in morphologic CR but had ≥ 5% aberrant blasts by MFC (range, 5.4% to 38%); six died within 2 years, with one patient alive at 58.6 months.

Relative Prognostic Effect of MFC-MRD After C1 and C2 by Genetic/Risk Score Subgroup

In AML17, patients received two courses of induction regardless of remission status after C1, but C2 differed for patients designated as poor risk by trial risk score. Analyses of survival and relapse by MFC-MRD status of patients with disease in CR/CRi for C1 (n = 1,010) and C2 (n = 803) were performed stratified by cytogenetic[28] and trial risk subgroups (Fig 2; Appendix Fig A4, online only) to investigate the relative prognostic effect from clearance of blasts below MFC-MRD detection threshold at either of these response assessment time points. There was some evidence that the benefit from MFC-MRD negativity on OS was lower in patients at poor risk compared with other subgroups with the NCRI AML17 treatment schedule (P for test for trend = .01 for C1; $P = .05$ for C2). Overall, MFC-MRD status appeared more prognostic for relapse and OS at C2 (relapse: HR, 1.88 [95% CI, 1.50 to 2.36], $P < .001$; survival: HR, 1.77 [95% CI, 1.41 to 2.22], $P < .001$) than C1 (relapse: HR, 1.70 [95% CI, 1.40 to 2.06], $P < .001$; survival: HR, 1.50 [95% CI, 1.23 to 1.84], $P < .001$), although this difference diminished when C1 analysis was restricted to patients who received C2 and survived at least 30 days post-C2 (relapse: HR, 1.80 [95% CI, 1.49 to 2.18], $P < .001$; survival:, HR, 1.87 [95% CI, 1.52 to 2.29], $P < .001$).

Outcomes of Combined C1 Response Status and C2 MFC-MRD Status

In patients with response/MFC-MRD data for both C1 and C2 time points (n = 693), C2 MFC-MRD positivity remained significant on OS and relapse when adjusting for C1 response (5-year survival: HR, 1.79 [95% CI, 1.38 to 2.32], $P < .001$; relapse: HR, 1.52 [95% CI, 1.18 to 1.96], $P = .001$; Fig 3). A total of 24 patients converted from C1 MRD− to C2 MRD+, with a particularly poor prognosis (n = 15 relapses; n = 13 deaths); one had adverse risk cytogenetics and five had *Flt3*-ITD mutations (Appendix Table A1). Patients who were MRD− at both C1 and C2 had the best outcome (n = 224; n = 76 relapses; n = 58 deaths); of these, 80.8% were at good or standard risk and 26.3% were *NPM1*-wt patients at standard risk (Appendix Table A2).

MRD Status Combined With Peripheral Count Recovery

We examined the additional prognostic effect of combining MRD status with response by peripheral count recovery post-C1 and C2 (Appendix Table A3). The frequencies of CRi as best response in the total cohort were similar in MRD+ versus MRD− patients post-C1 (9.3% *v* 9.6%) and C2 (13.1% *v* 12.0%); CRi frequencies were not relatively increased in the *NPM1*-wt standard-risk subgroup. C1 CRi was associated with significantly decreased 5-year OS for total (39% *v* 53%; $P = .002$) and in MRD+ (19% *v* 45%; $P = .001$), but not for MRD− patients. MRD+ *NPM1*-wt patients at standard risk in CRi also had a lower OS at 5 years (25% *v* 48%; $P = .4$), although difference was not significant. The effect of CRi versus CR was smaller post-C2, although outcomes were still worse in CRi/MRD+ patients. The reduced survival associated with CRi was not due to increased relapse.

Outcome by MFC-MRD Status for NPM1-wt Patients at Standard Risk

Because it is possible that the most appropriate MFC-MRD cutoff level for discriminating outcome may differ among AML genetic subgroups, we compared the 5-year cumulative incidence of relapse for C1 MRD− versus MRD+ < 0.1% versus MRD+ ≥ 0.1% by our assay in core binding factor (CBF)-AML and *NPM1*-mutated as well as *NPM1*wt standard-risk patients. For patients with CBF-AML and *NPM1* mutation, post-C1 MRD+ at any level (< 0.1% or ≥ 0.1%) significantly increased relapse (Appendix Fig A5, online only). However, in the *NPM1*-wt standard-risk subgroup, low-level MRD+ (< 0.1%) post-C1 did not alter relapse risk compared with MRD− but was associated with a higher cumulative incidence of relapse (CIR) when detected post-C2 (Fig 4A). MRD+ levels of ≥ 0.1% detected in 35% and 13% *NPM1*-wt patients at standard risk post-C1 and post-C2, respectively, predicted a high probability of relapse (C1 3-year CIR, 68%; C2 CIR, 89%). MRD status after second induction was also significantly prognostic for survival: 33% for any level of MRD positivity versus 63% for MRD− at 5 years (3 years, 47% *v* 69%; $P = .003$; Fig 4B).

Of the 204 *NPM1*-wt patients at standard risk who had C2 MRD data, 83 had an allograft (n = 44 in first CR: n = 29 MRD− and n = 15 MRD+). When survival was censored at any SCT, rates of 5-year OS were 35% versus 88% (3 years, 47% *v* 88%; $P < .001$; Appendix Fig A6, online only).

We next investigated the effect of SCT in first CR according to C2 MRD status in Mantel-Byar analyses. Although numbers were small, results suggested that transplant might be considered in MRD+ (HR, 0.72; 5% CI, 0.31 to 1.69) but not MRD− patients (HR, 1.68 [95% CI, 0.75 to 3.85]; *P* for interaction = .16; Fig 4C).

DISCUSSION

Response to induction therapy is a powerful prognostic indicator in AML. There are, however, differing practices for the implementation of technologies that measure residual leukemia to assess response. Flow cytometry is often used to support the definition of CR by morphology; those centers with access to experienced laboratories, including some trial groups, have extended its use to define CR without MRD.[5] It has been reported that outcomes after myeloablative SCT for patients with pretransplant MFC-MRD < 5% resemble those with at least 5% blasts by morphology.[13] This and the similar event-free survival observed in approximately 80 pediatric patients with MRD positivity after first induction, whether < 5% or ≥ 5% blasts by morphology,[7] suggest that dichotomizing patients by a 5% blast CR cutoff fails to capture some prognostic information. Our results confirm this. By incorporating MFC-MRD with established response criteria of PR and RD, distinct prognostic groups for 5-year survival emerge after the first

Table 2. Outcomes by Response After Course 1

Patient subgroups	PR	MRD+ in CR	Unadjusted HR/OR (95% CI); *P*	Adjusted* HR/OR, (95% CI); *P*	RD	PR or MRD+ in CR	Unadjusted HR/OR (95% CI); *P*	Adjusted* HR/OR (95% CI); *P*
No. of total patients	197	577			223	774		
OS	35 (41)	44 (51)	1.50 (1.18 to 1.91); .001	1.32 (1.05 to 1.66); .02	24 (30)	41 (48)	2.43 (1.93 to 3.06); < .001	1.61 (1.31 to 1.97); < .001
OS censored at SCT	41 (48)	49 (54)	1.81 (1.25 to 2.61); .002	1.62 (1.16 to 2.25); .004	11 (11)	48 (53)	8.17 (5.76 to 11.6); < .001	2.64 (2.02 to 3.45); < .001
No. of patients at good to standard risk	53	352			40	405		
OS	49 (58)	51 (59)	1.11 (.69 to 1.76); .7	1.320 (.65 to 2.68); .4	33 (33)	51 (59)	16.6 (9.24 to 30.0); < .001	2.28 (1.38 to 3.75); < .001
OS censored at SCT	67 (73)	60 (66)	1.05 (.54 to 2.06); .9	1.38 (.68 to 2.80); .4	37 (at 3, 4 years)	61 (67)	22.4 (8.23 to 60.8); < .001	3.56 (1.78 to 7.12); < .001
No. of NPM1 WT patients at standard risk	27	149			34	176		
OS	38 (55)	47 (55)	1.10 (.60 to 2.04); .8	1.18 (.61 to 2.28); .6	28 (28)	45 (55)	3.09 (1.65 to 5.80); < .001	2.13 (1.21 to 3.75); .008
OS censored at SCT	53 (71)	52 (59)	1.38 (.52 to 3.63); .5	1.51 (.58 to 3.93); .4	28 (at 3 years)	53 (60)	18.2 (6.38 to 52.1); < .001	3.88 (1.68 to 8.94); .001

NOTE. Data reported as 5-year (3-year) percentages unless otherwise indicated.
Abbreviations: CR, complete remission; CRi, complete remission with absolute neutrophil count < 1,000/μL or thrombocytopenia < 100,000/μL; HR, hazard ratio; MRD, measurable residual disease; OR, odds ratio; OS, overall survival; PR, partial response; SCT, stem-cell transplant; WT, wild type.
*Adjusted analyses included age, WBC count, sex, performance status, disease type (secondary or de novo), and cytogenetic group.

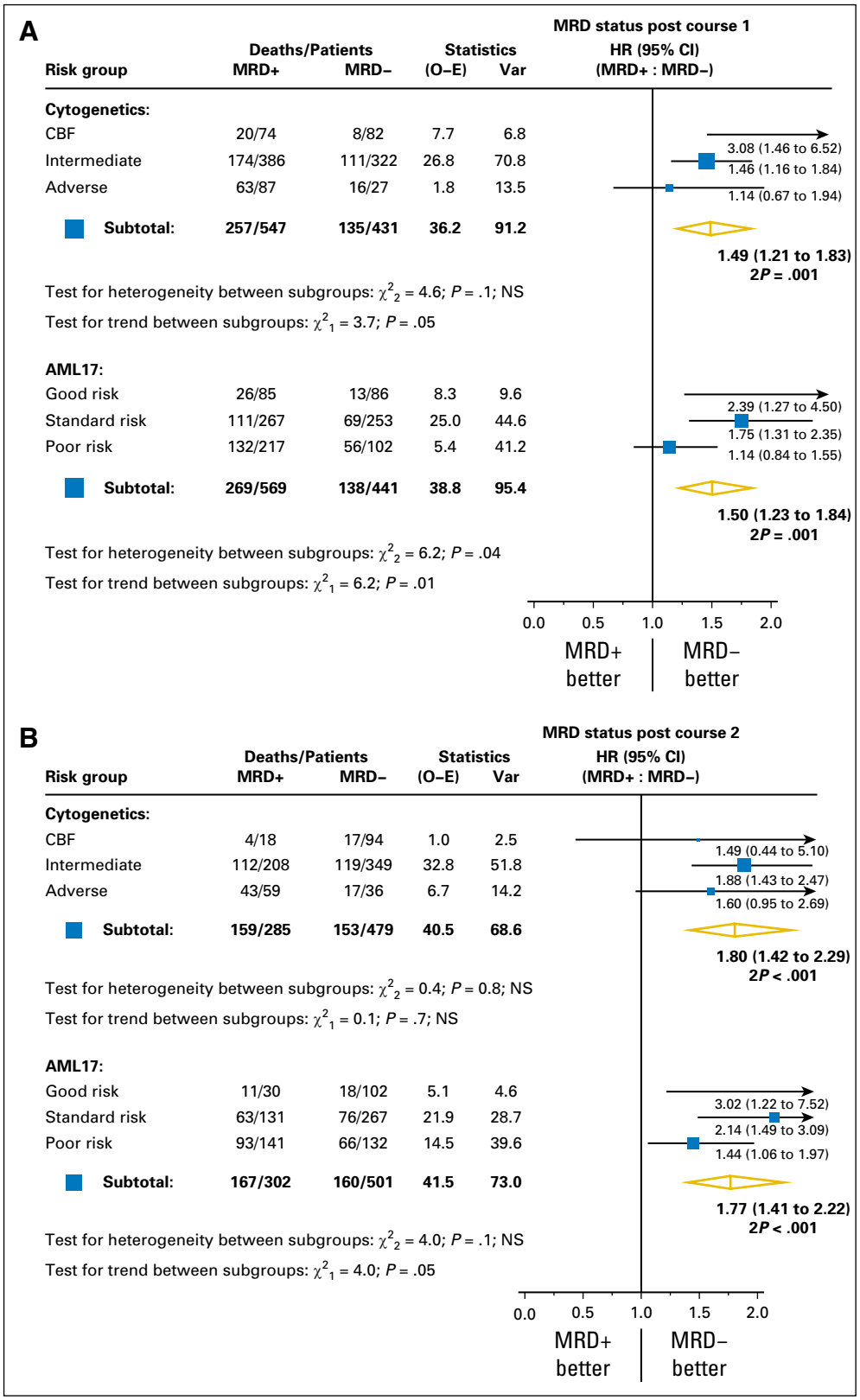

Fig 2. Forest plots for overall survival by multiparametric flow cytometry-MRD status for patients in complete remission. (A) After course 1. (B) After course 2, stratified by cytogenetic risk group and NCRI AML17 risk score group. CBF, core binding factor; HR, hazard ratio; MRD, measurable residual disease; NS, not significant; O-E, observed minus expected; Var, variance.

course of standard induction. Importantly, the response subgroup with intermediate outcome comprises patients on either side of the current CR blast threshold, those with MRD positivity in CR and those who are refractory but clinically classified as a PR; both responses are associated with similar 5-year survival, particularly in patients otherwise allocated as belonging to good- or standard-risk subgroups. This is also the case when PR is defined by European LeukemiaNet criteria[5,21](Appendix Fig A7, online only). From this,

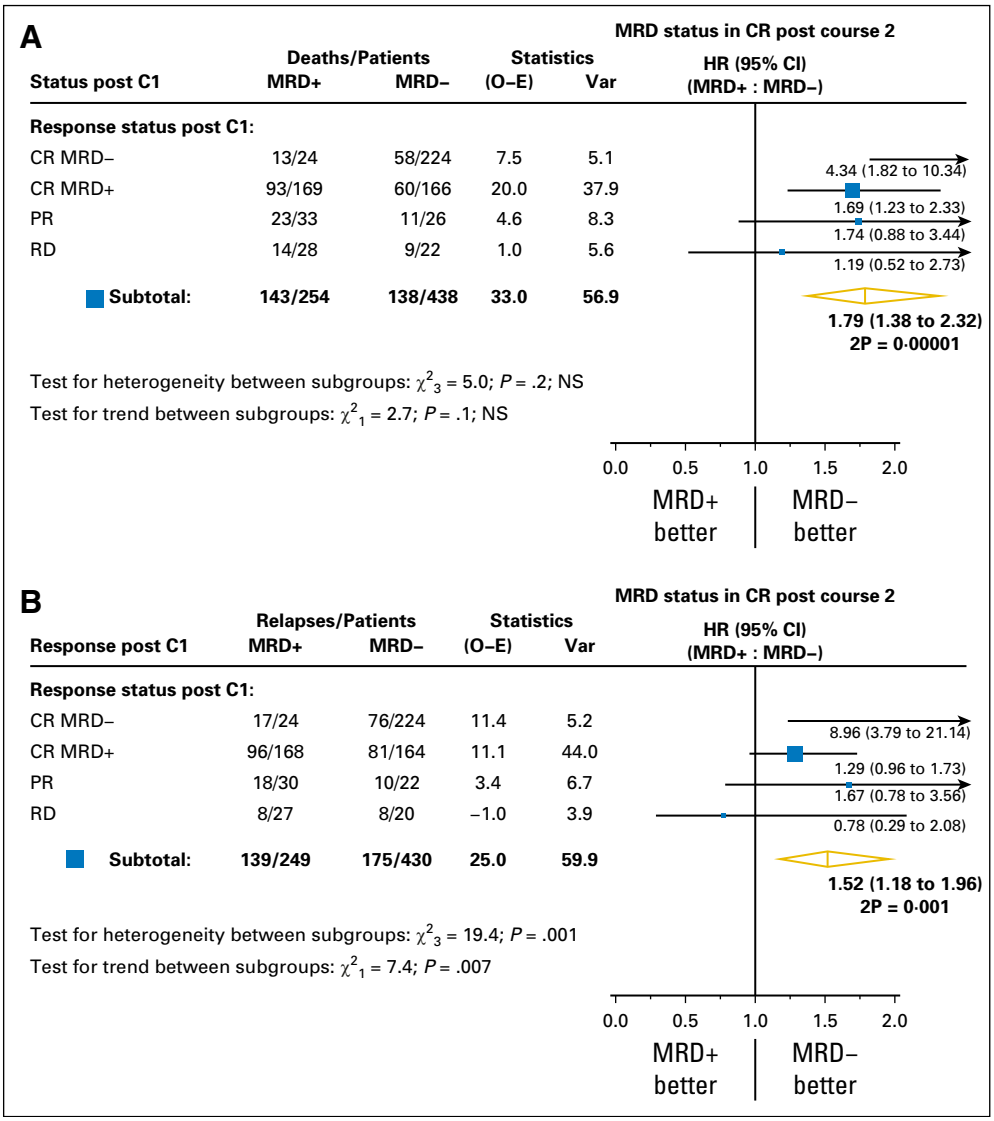

Fig 3. Forest plots for (A) overall survival and (B) relapse by combined response data after courses 1 and 2. Effect of multiparametric flow cytometry-MRD status in CR after course 2 stratified by post-C1 response status. C1, course 1; CR, complete remission; HR, hazard ratio; MRD, measurable residual disease; NS, not significant; O-E, observed minus expected; PR, partial remission; RD, resistant disease; Var, variance.

three post-C1 response categories could be proposed: RD, PR (MFC-MRD+ whether below or above 5% blast threshold), and CR/CRi without MRD. CRi was an independent risk factor to MRD in a study that included patients with relapsed or refractory AML and differing induction intensities.[29,30] From our data, outcomes for patients newly diagnosed with AML achieving negative MRD are equivalent between CRi and CR after a single standard induction. However, the relatively few patients in our cohort (4.8%) with both CRi and MRD positivity after C1 had as poor survival (OS, 19% for all;, 25% for *NPM1*-wt patients at standard risk) as patients with RD.

For those completing a second induction with a CR/CRi, MRD status after C2 increased prognostic discrimination. Although sample attrition bias may limit analyses comparing time points, MRD negativity post-C2 improved outcome overall even when adjusting for slower blast clearance by C1 response. This differs from our previous results in older adults[17] and might reflect the better treatment tolerance and mutation profiles of younger adults. However, after the second daunorubicin/cytosine arabinoside induction, approximately 33% of patients at standard risk and approximately 34% of *NPM1*-wt patients at standard risk in CR/CRi had persistent BM MRD by our assay. Whether detectable MFC-MRD after completion of conventional induction is a sufficiently specific prognostic surrogate to guide therapy has been debated. The postconsolidation time point was more informative in the GIMEMA (Gruppo Italiano Malattie Ematologiche dell'Adulto) study for a cohort of which approximately 70% had intermediate cytogenetics.[31,32] This suggests that in a proportion of those with postinduction MRD positivity, consolidation may confer a favorable outcome by additional MFC-MRD clearance (although it is of note that for some younger adults in the GIMEMA trials, the induction/consolidation regimen comprised two courses in total). Genetic profile, treatment intensity, and the later effects of any transplant may also modify interpretation and utility of MFC-MRD to inform postremission therapy. Our data are consistent with this because the prognostic effect as well as best MFC-MRD cutoff level differed between AML risk groups; MRD status appeared less discriminatory in the patients at poor risk. Importantly, however, in the *NPM1*-wt standard-risk subgroup,

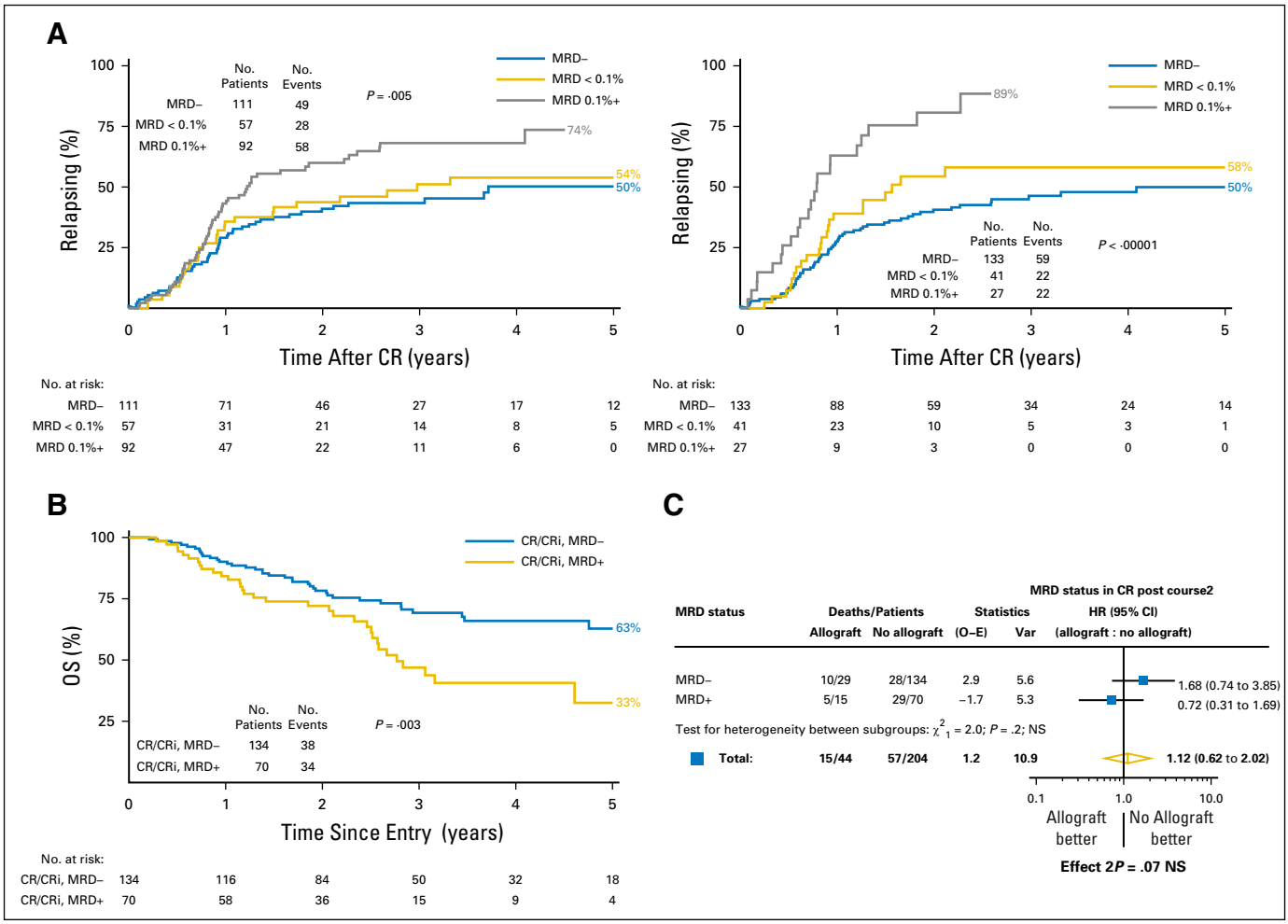

Fig 4. Standard-risk *NPM1*-wild type. (A) Cumulative incidence of relapse by MRD level. (MRD− *v* MRD+ < 0.1% *v* MRD+ ≥ 0.1%) after courses 1 and 2. (B) Overall survival (OS) according to MRD status after course 2 (MRD− *v* MRD+). (Not shown: MRD+ ≥ 0.1%, OS of 24%; MRD+ < 0.1%, OS of 39%). (C) Mantel-Byar analysis for survival according to first CR stem-cell transplant by MRD after course 2. CR, complete remission; CRi, complete remission with absolute neutrophil count < 1,000/μL or thrombocytopenia < 100,000/μL; HR, hazard ratio; MRD, measurable residual disease; NS, not significant; O-E, observed minus expected; Var, variance.

detectable MFC-MRD at ≥ 0.1% early in treatment was associated with significantly higher relapse rates (89% after C2). The false-negative 50% CIR observed for postinduction MFC-MRD negative patients in the *NPM1*-wt, standard-risk subgroup could reflect MFC-MRD sensitivity limitations, although a similar CIR was observed for patients with *DNTM3A*/*NPM1* mutations who were MRD negative by *NPM1*-mutated transcript RT-qPCR.[23] Exploratory analyses could not identify any significant clinical parameters that predicted MRD− relapses. Longitudinal, broad molecular studies may disclose whether increased preleukemic instability reinitiating AML[33,34] or persistence of pretreatment minor or major leukemic clones[35,36] contributes to these false-negative relapse risks. Notwithstanding, *NPM1*-wt patients at standard risk who achieved MRD negativity post-C2 had significantly better survival rates. Because their survival rate increased to 88% when censored for transplant, there is the possibility that transplant in first remission could be avoided in this subset. The Mantel-Byar analysis supports this with some evidence of interaction, although this should be interpreted cautiously because of the small number of patients and the interaction was not significant.

Transplant decisions have mainly been arbitrary in this subgroup, with no accepted approach to distinguish those patients likely to be cured with chemotherapy alone (or those whose response is likely to be successful after salvage therapy if they do relapse) from those who benefit from transplantation in first remission or potentially experimental therapy. Our results suggest that allogeneic transplant in first remission could be directed to those who are MRD+ rather than MRD−. This is the first indication that MRD status might have utility in directing therapy for *NPM1*-wt patients at standard risk despite their molecular heterogeneity. Large patient data sets likely requiring collaborative efforts will determine whether integrating MFC-MRD status with genomic profiles[37,38] further informs outcome prediction.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at jco.org.

AUTHOR CONTRIBUTIONS

Conception and design: Sylvie D. Freeman, Robert K. Hills, David Grimwade, Alan K. Burnett, Nigel H. Russell
Provision of study materials or patients: Asim Khwaja
Collection and assembly of data: Sylvie D. Freeman, Robert K. Hills, Paul Virgo, Naeem Khan, Steve Couzens, Richard Dillon, Amanda Gilkes, Laura Upton, Ove Juul Nielsen, James D. Cavenagh, Gail Jones, Asim Khwaja, Paul Cahalin, Ian Thomas
Data analysis and interpretation: Sylvie D. Freeman, Robert K. Hills, Alan K. Burnett, Nigel H. Russell
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

REFERENCES

1. Wheatley K, Burnett AK, Goldstone AH, et al: A simple, robust, validated and highly predictive index for the determination of risk-directed therapy in acute myeloid leukaemia derived from the MRC AML 10 trial. Br J Haematol 107:69-79, 1999

2. Schlenk RF, Benner A, Hartmann F, et al: Risk-adapted postremission therapy in acute myeloid leukemia: Results of the German multicenter AML HD93 treatment trial. Leukemia 17: 1521-1528, 2003

3. Kern W, Haferlach T, Schoch C, et al: Early blast clearance by remission induction therapy is a major independent prognostic factor for both achievement of complete remission and long-term outcome in acute myeloid leukemia: Data from the German AML Cooperative Group (AMLCG) 1992 Trial. Blood 101:64-70, 2003

4. O'Donnell MR, Tallman MS, Abboud CN, et al: Acute myeloid leukemia, version 3.2017, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 15:926-957, 2017

5. Döhner H, Estey E, Grimwade D, et al: Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood 129:424-447, 2017

6. Bassan R: Using minimal residual disease to improve treatment response definitions and hematopoietic cell transplantation strategy in acute leukemia. J Clin Oncol 34:300-302, 2016

7. Inaba H, Coustan-Smith E, Cao X, et al: Comparative analysis of different approaches to measure treatment response in acute myeloid leukemia. J Clin Oncol 30:3625-3632, 2012

8. Loken MR, Alonzo TA, Pardo L, et al: Residual disease detected by multidimensional flow cytometry signifies high relapse risk in patients with de novo acute myeloid leukemia: A report from Children's Oncology Group. Blood 120:1581-1588, 2012

9. Ouyang J, Goswami M, Tang G, et al: The clinical significance of negative flow cytometry immunophenotypic results in a morphologically scored positive bone marrow in patients following treatment for acute myeloid leukemia. Am J Hematol 90:504-510, 2015

10. Buccisano F, Maurillo L, Del Principe MI, et al: Prognostic and therapeutic implications of minimal residual disease detection in acute myeloid leukemia. Blood 119:332-341, 2012

11. Grimwade D, Freeman SD: Defining minimal residual disease in acute myeloid leukemia: Which platforms are ready for "prime time"? Hematology (Am Soc Hematol Educ Program) 2014:222-233, 2014

12. Ossenkoppele G, Schuurhuis GJ: MRD in AML: Does it already guide therapy decision-making? Hematology (Am Soc Hematol Educ Program) 2016: 356-365, 2016

13. Araki D, Wood BL, Othus M, et al: Allogeneic hematopoietic cell transplantation for acute myeloid leukemia: Time to move toward a minimal residual disease-based definition of complete remission? J Clin Oncol 34:329-336, 2016

14. Hourigan CS, Goswami M, Battiwalla M, et al: When the minimal becomes measurable. J Clin Oncol 34:2557-2558, 2016

15. Hourigan CS, Gale RP, Gormley NJ, et al: Measurable residual disease testing in acute myeloid leukaemia. Leukemia 31:1482-1490, 2017

16. Buccisano F, Maurillo L, Spagnoli A, et al: Cytogenetic and molecular diagnostic characterization combined to postconsolidation minimal residual disease assessment by flow cytometry improves risk stratification in adult acute myeloid leukemia. Blood 116:2295-2303, 2010

17. Freeman SD, Virgo P, Couzens S, et al: Prognostic relevance of treatment response measured by flow cytometric residual disease detection in older patients with acute myeloid leukemia. J Clin Oncol 31:4123-4131, 2013

18. 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 30: 2080-2083, 2016

19. Ravandi F, Jorgensen J, Borthakur G, et al: Persistence of minimal residual disease assessed by multiparameter flow cytometry is highly prognostic in younger patients with acute myeloid leukemia. Cancer 123:426-435, 2017

20. Terwijn M, van Putten WL, 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 31:3889-3897, 2013

21. Cheson BD, Bennett JM, Kopecky KJ, et al: Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. J Clin Oncol 21:4642-4649, 2003 [Erratum: J Clin Oncol 20041;22(3):576

22. Ferguson P, Hills RK, Grech A, et al: An operational definition of primary refractory acute myeloid leukemia allowing early identification of patients who may benefit from allogeneic stem cell transplantation. Haematologica 101:1351-1358, 2016

23. Ivey A, Hills RK, Simpson MA, et al: Assessment of minimal residual disease in standard-risk AML. N Engl J Med 374:422-433, 2016

24. 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 35:185-193, 2017

25. Burnett AK, Hills RK, Wheatley K, et al: A sensitive risk score for directing treatment in younger patients with AML. Blood 108:18, 2006

26. Ling V, Burnett AK, Bradstock K, et al: Utility of a clinical risk score to identify high-risk patients with de novo acute myeloid leukaemia in first remission after high-dose cytarabine (HiDAC) based induction chemotherapy. Br J Haematol 160: 861-863, 2013

27. Early Breast Cancer Trialists' Cooperative Group . Treatment of Early Breast Cancer. 1. Worldwide Evidence 1985-1990. Oxford, UK, Oxford University Press, 1990

28. 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 116:354-365, 2010

29. Chen X, Xie H, Estey EH: Reply to D. Przepiorka et al. J Clin Oncol 33:3676-3677, 2015

30. Chen X, Xie H, Wood BL, et al: Relation of clinical response and minimal residual disease and their prognostic impact on outcome in acute myeloid leukemia. J Clin Oncol 33:1258-1264, 2015

31. Buccisano F, Maurillo L, Gattei V, et al: The kinetics of reduction of minimal residual disease impacts on duration of response and survival of patients with acute myeloid leukemia. Leukemia 20: 1783-1789, 2006

32. Maurillo L, Buccisano F, Del Principe MI, et al: Toward optimization of postremission therapy for residual disease-positive patients with acute myeloid leukemia. J Clin Oncol 26:4944-4951, 2008

33. da Silva-Coelho P, Kroeze LI, Yoshida K, et al: Clonal evolution in myelodysplastic syndromes. Nat Commun 8:15099, 2017

34. Makishima H, Yoshizato T, Yoshida K, et al: Dynamics of clonal evolution in myelodysplastic syndromes. Nat Genet 49:204-212, 2017

35. Parkin B, Londoño-Joshi A, Kang Q, et al: Ultrasensitive mutation detection identifies rare residual cells causing acute myelogenous leukemia relapse. J Clin Invest 127:3484-3495, 2017

36. Shlush LI, Mitchell A, Heisler L, et al: Tracing the origins of relapse in acute myeloid leukaemia to stem cells. Nature 547:104-108, 2017

37. Bullinger L, Döhner K, Döhner H: Genomics of acute myeloid leukemia diagnosis and pathways. J Clin Oncol 35:934-946, 2017

38. Gerstung M, Papaemmanuil E, Martincorena I, et al: Precision oncology for acute myeloid leukemia using a knowledge bank approach. Nat Genet 49: 332-340, 2017

Affiliations

Sylvie D. Freeman and **Naeem Khan**, University of Birmingham, Birmingham; **Robert K. Hills**, **Amanda Gilkes**, **Laura Upton**, **Ian Thomas**, and **Alan K. Burnett**, Cardiff University; **Steve Couzens** University Hospital of Wales, Cardiff; **Paul Virgo**, North Bristol NHS Trust, Bristol; **Richard Dillon** and **David Grimwade**, King's College London School of Medicine; **James D. Cavenagh**, Queen Mary University of London; **Asim Khwaja**, University College London, London; **Gail Jones**, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle; **Paul Cahalin**, Blackpool Teaching Hospital NHS Foundation Trust, Blackpool; **Nigel H. Russell**, Nottingham University Hospital, Nottingham, United Kingdom; **Ove Juul Nielsen**, Rigshospitalet, Copenhagen, Denmark.

Support

Supported by Cancer Research UK and National Institutes of Health Research, which contributed nonprofit funding for this manuscript; no commercial funding was received.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Measurable Residual Disease at Induction Redefines Partial Response in Acute Myeloid Leukemia and Stratifies Outcomes in Patients at Standard Risk Without *NPM1* Mutations

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/site/ifc.

Sylvie D. Freeman
No relationship to disclose

Robert K. Hills
No relationship to disclose

Paul Virgo
No relationship to disclose

Naeem Khan
No relationship to disclose

Steve Couzens
No relationship to disclose

Richard Dillon
No relationship to disclose

Amanda Gilkes
No relationship to disclose

Laura Upton
No relationship to disclose

Ove Juul Nielsen
Consulting or Advisory Role: Roche
Travel, Accommodations, Expenses: Roche

James D. Cavenagh
No relationship to disclose

Gail Jones
No relationship to disclose

Asim Khwaja
Consulting or Advisory Role: Jazz Pharmaceuticals, Karus Therapeutics

Paul Cahalin
No relationship to disclose

Ian Thomas
No relationship to disclose

David Grimwade
No relationship to disclose

Alan K. Burnett
Employment: CTI Life Sciences
Consulting or Advisory Role: Jazz Pharmaceuticals

Nigel H. Russell
No relationship to disclose

Acknowledgment

We thank the staff of the Haematology Trials Office, Cardiff Experimental Cancer Medicine Centre and reference laboratories for supporting the trial; members of the NCRI AML working group and to our other coinvestigators, research nurses, and participants who provided samples from the NCRI AML Trial centers (Data Supplement).

Appendix

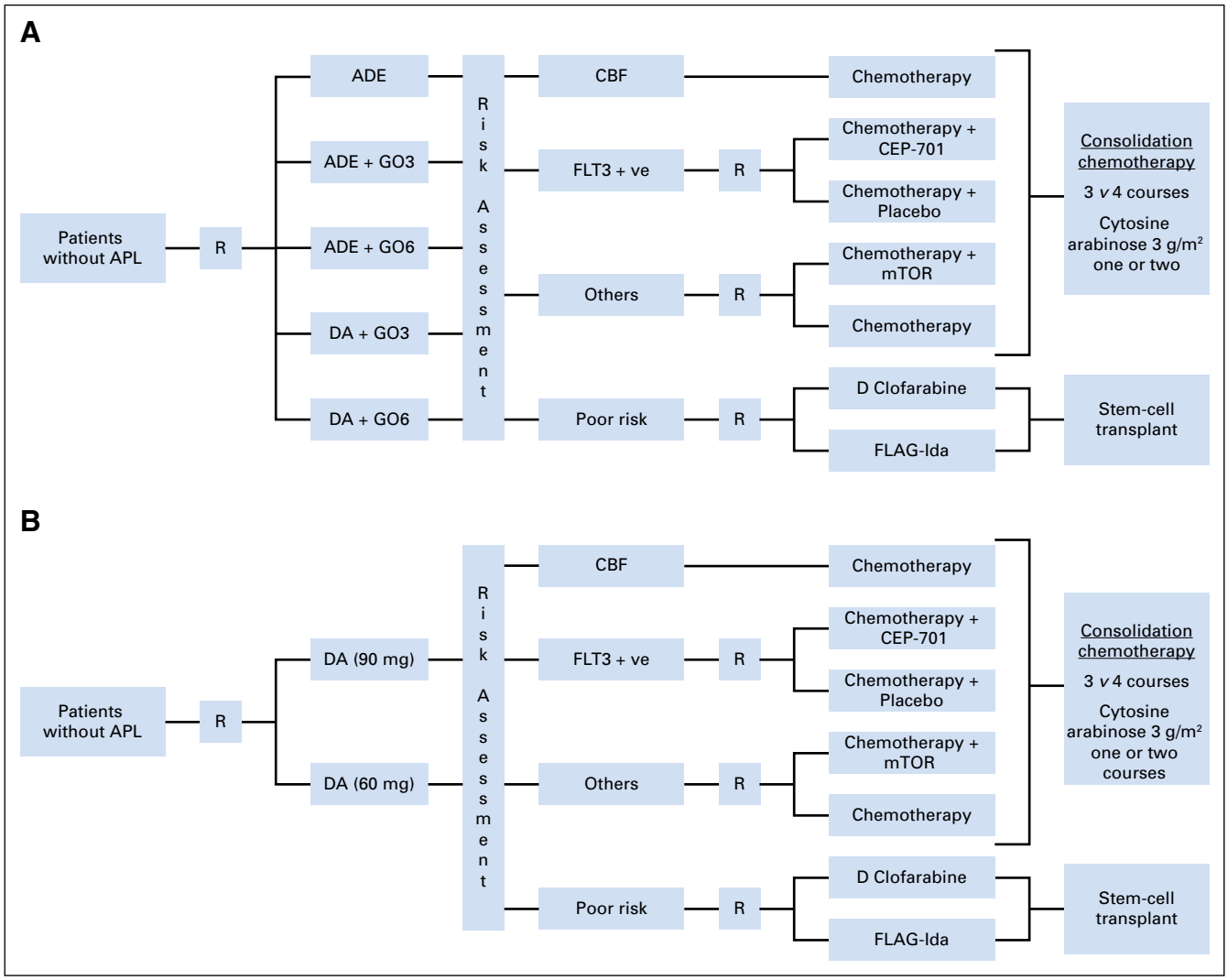

Fig A1. Flowchart of treatments given to patients in the NCRI AML17 trial. (A) Pre-October 2011 (induction gemtuzumab ozogamicin randomization). (B) Post-October 2011 (daunorubicin dose randomization in induction). Note: patients who did not satisfy the hepatic entry criteria (liver function $< 2 \times$ ULN) in (A) were allocated ADE; until June, 2010 the consolidation randomization was MACE vs MACE/MidAC; the DA dose randomization was closed in October, 2013, and patients subsequently received DA (60 mg); the lestaurtinib (CEP-701) randomization closed in October 2012; the mTOR (everolimus) randomization closed in August, 2012; the high risk randomization in October, 2012. All core binding factor (CBF) leukemias were eligible for gemtuzumab ozogamicin and were given 3 mg/m^2 with course 2 if they did not receive it by gemtuzumab ozogamicin randomization with course 1. From June, 2012 patients with informative real-time quantitative polymerase chain reaction (RT-qPCR) MRD markers could enter the 'Monitor vs no Monitor' randomization that investigates the impact of serial RT-qPCR monitoring post completion of treatment on outcome, quality of life and health economics. ADE, cytarabine, daunorubicin, and etoposide; APL, acute promyelocytic leukemia; CBF, core binding factor; CEP-701, lestaurtinib; DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin (3 or 6 mg/m^2); FLAG-Ida, fludarabine, cytarabine, GCSF, and idarubicin; FLT3, FMS-like tyrosine kinase-3; mTOR, everolimus; R, randomization.

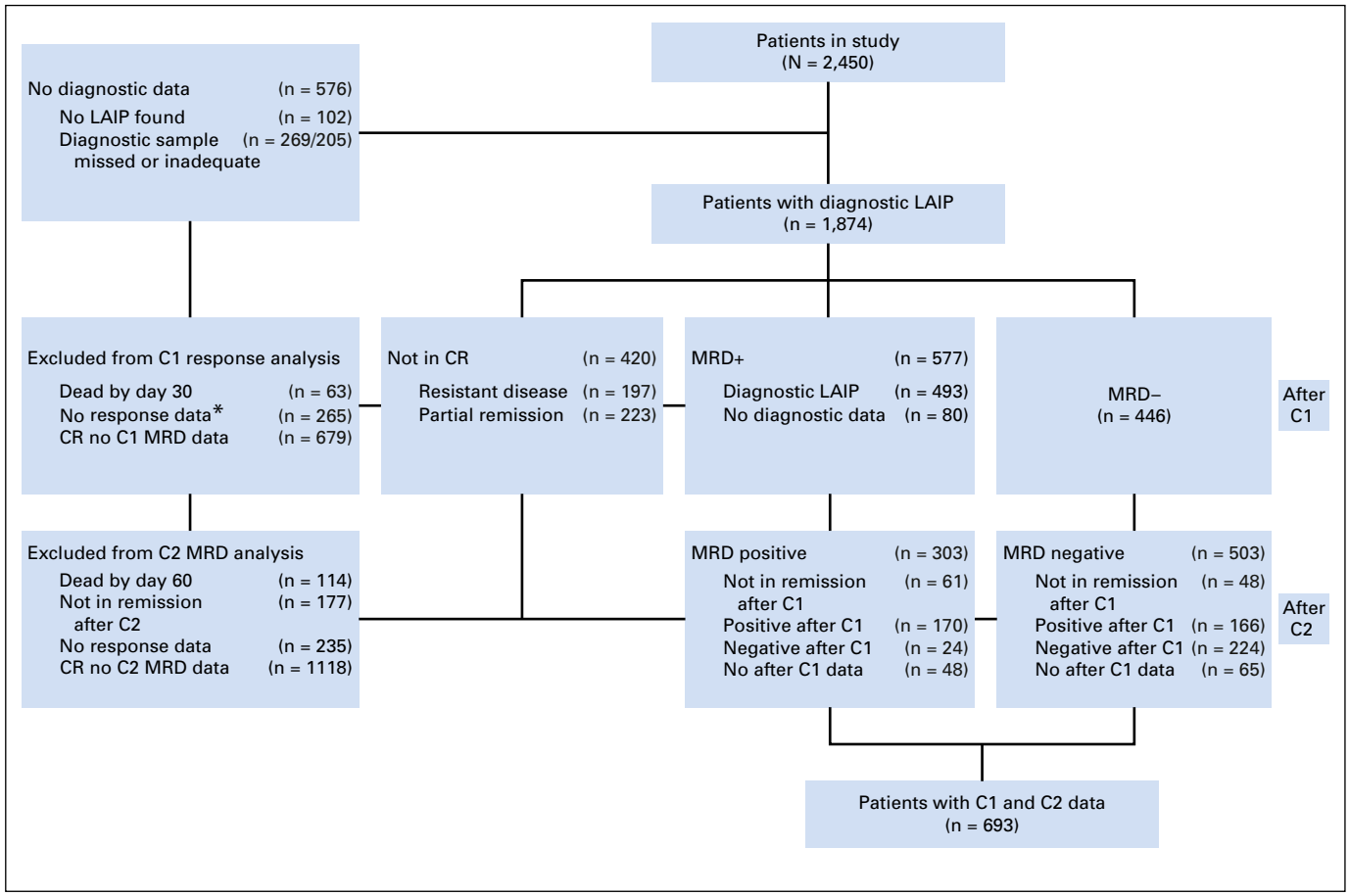

Fig A2. CONSORT diagram. Outline of patient sample flow for MRD study. (*) Includes patients for whom remission status could not be classified as exact timing of any remission was unavailable. CR, complete remission; C1, course 1, C2, course 2. LAIP, leukemia-associated–immunophenotype.

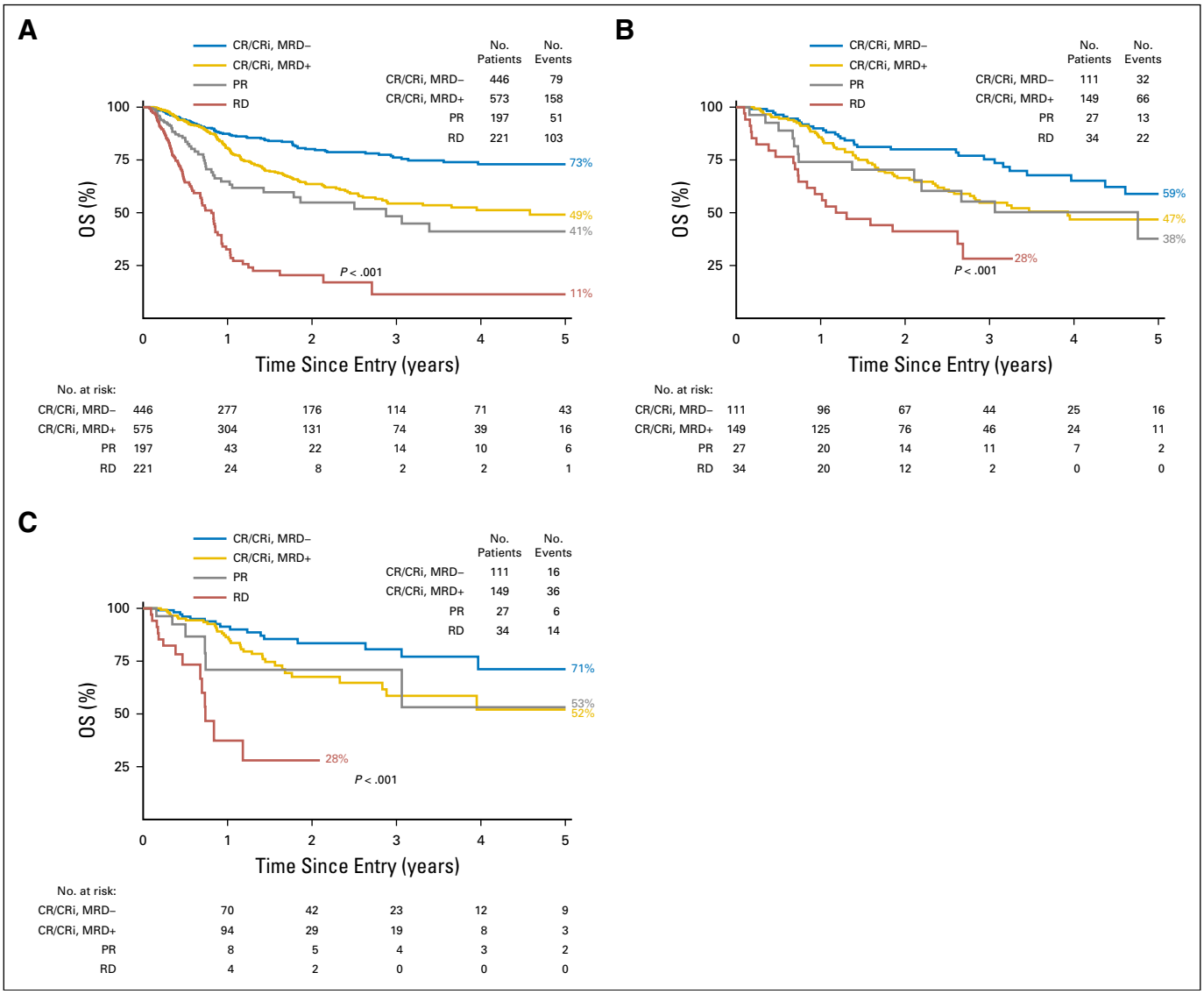

Fig A3. OS according to response status after course 1. (A) All patients. OS censored at allogeneic SCT. (B) *NPM1*–wild-type patients at standard risk. (C) *NPM1*–wild-type patients at standard risk, censored at allogeneic SCT. CR, complete remission; MRD, measurable residual disease; OS, overall survival; PR, partial remission; RD, resistant disease; SCT, stem-cell transplantation.

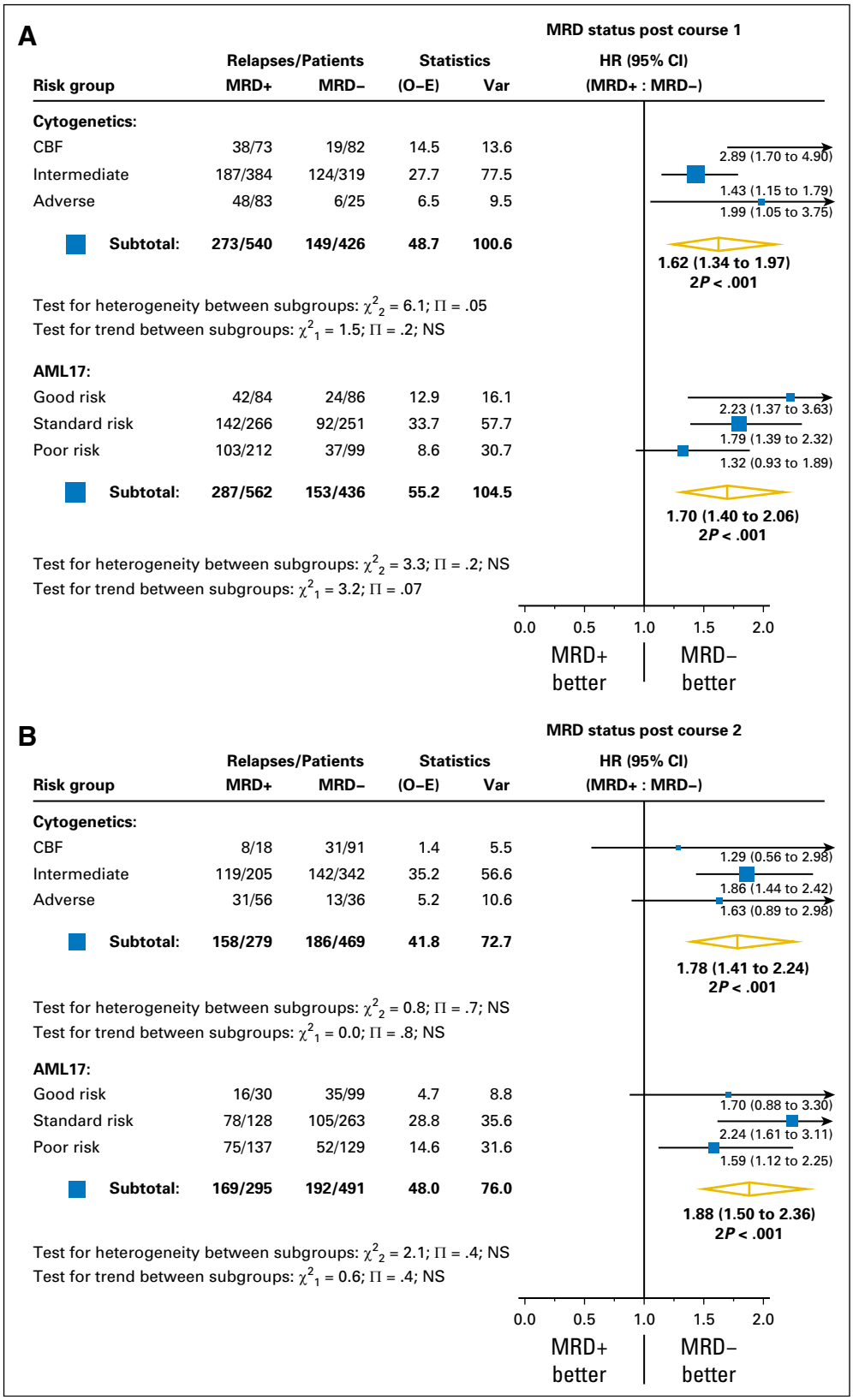

Fig A4. Forest plots for relapse by multiparametric flow cytometry-MRD status for patients in CR (A) after course 1 and (B) after course 2 stratified by cytogenetic risk group and NCRI AML 17 risk score group. CR, complete remission; MRD, measurable residual disease.

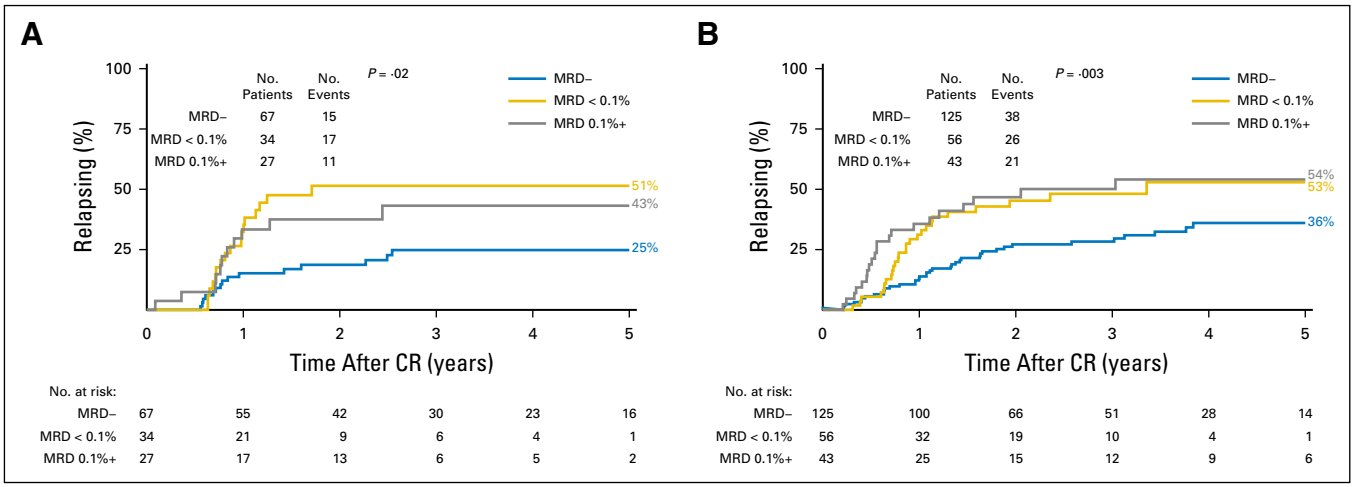

Fig A5. Cumulative incidence of relapse by multiparametric flow cytometry -MRD level. (MRD− *v* MRD+ < 0.1% *v* MRD+ ≥ 0.1%) after course 1. (A) CBF AML. (B) Standard-risk NPM1 mutant. AML, acute myeloid leukemia; CBF, core binding factor; MRD, measurable residual disease; MRD < 0.1%, MRD+ < 0.1%; MRD 0.1%+, MRD+ ≥ 0.1%.

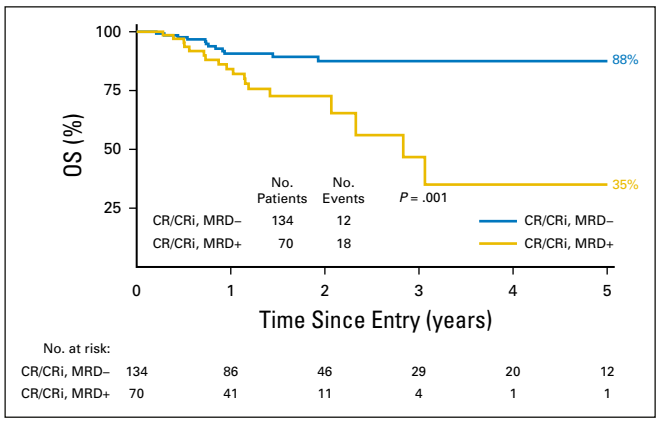

Fig A6. Standard-risk *NPM1*-wild type. Overall survival (OS) according to multiparametric flow cytometry-MRD status after course 2, censored at any allogeneic stem-cell transplantation. CR, complete remission; MRD, measurable residual disease.

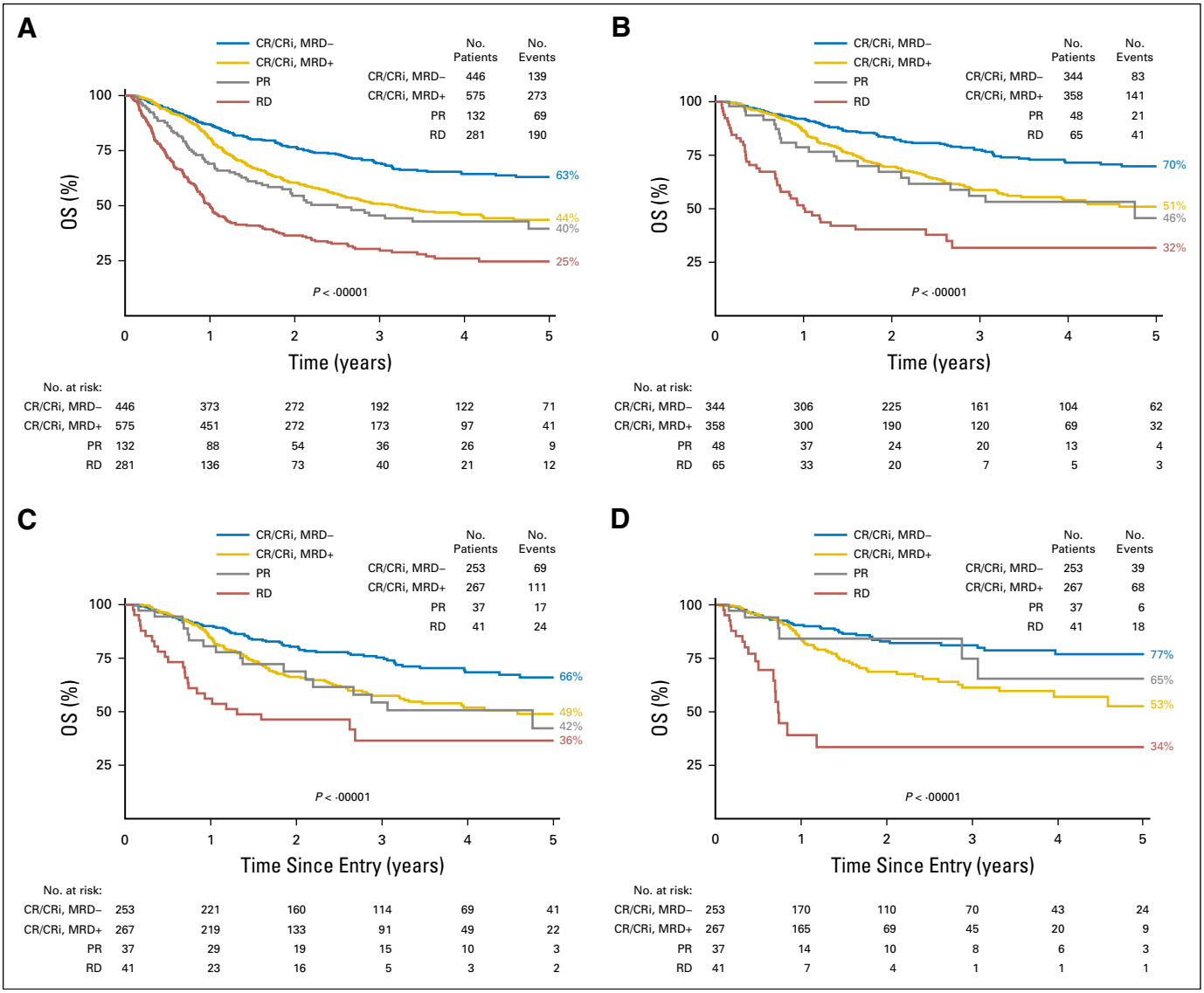

Fig A7. OS according to response status after course 1, applying European LeukemiaNet (ELN)/Cheson criteria for PR and RD instead of MRC criteria. (ELN criteria for PR: all hematologic criteria of CR; decrease of bone marrow blast percentage to 5% to 25% with decrease of pretreatment bone marrow blast percentage by ≥ 50%. (A) All patients. (B) Patients at good and standard risk (patients known to be at poor risk excluded). (C) Patients at standard risk. (D) Patients at standard risk, OS censored at allogeneic SCT. CR, complete remission; MRD, measurable residual disease; OS, overall survival; PR, partial remission; RD, resistant disease; SCT, stem-cell transplantation.

Table A1. Genetic Characteristics of C1 MFC-MRD- / C2 MFC-MRD+ Patients

Karyotype	Cytogenetic Risk Group	NPM1 Mutation	Flt3 ITD
46,XY,t(8;21)(q22;q22)[7]/47,idem,+8[3]	Good	Negative	Negative
46,XY[20]	Intermediate	NA	NA
46,XY[20]	Intermediate	Positive	Positive
46,XY[20]	Intermediate	Positive	Positive
46,XX[16]	Intermediate	Positive	Positive
46,XX[20]	Intermediate	Negative	Positive
46,XY[20]	Intermediate	Negative	Positive
46,XX[20]	Intermediate	Positive	Negative
46,XY[20]	Intermediate	Negative	Negative
46,XX[20]	Intermediate	Negative	Negative
46,XX[20]	Intermediate	Negative	Negative
46,XX[20]	Intermediate	Negative	Negative
46,XX,t(11;19)(q23;p13.1)[10]	Intermediate	Negative	Negative
46,XY,del(9)(q?2q?3)[9]/46,XY[2]	Intermediate	Negative	Negative
46,XX,t(2;9)(p22;p21)[12]/46,XX[1]	Intermediate	Negative	Negative
46,XY,t(6;9)(p23;q24)[9]/46,XY[1]	Intermediate	Negative	Negative
46,XY[20]	Intermediate	Negative	Negative
45,XX,dic(17;18)(p11.2;p11.2)[9]/46,XX[1]	Intermediate	Negative	Negative
46,XX[20]	Intermediate	Negative	Negative
46,XY[20]	Intermediate	Negative	Negative
47,XY,+8[6]/ 46,XY[4]	Intermediate	Negative	Negative
50,XY,+X,+4,t(10;11)(p12;q14),+15,+19[9]/ 46,XY[1]	Adverse	Negative	Negative
Failed	NA	Negative	Negative
Failed	NA	Negative	Negative

Abbreviations: ITD, internal tandem duplication; NA, not applicable.

Table A2. Correlation of Risk Group and Clinical Response by Combined Course 1 and 2 Response Status

Risk Group	Total	C1 MRD−/ C2 MRD−	C1 MRD−/ C2 MRD+	C1 MRD+/ C2 MRD−	C1 MRD+/ C2 MRD+	C1 PR/ C2 MRD−	C1 PR/ C2 MRD+	C1 RD/ C2 MRD−	C1 RD/ C2 MRD+
All (patients with both C1 and C2 data)	693	224 (32.3)	24 (3.5)	166 (24.0)	170 (24.5)	26 (3.8)	33 (4.8)	22 (3.2)	28 (4.0)
Post-C1 risk score									
Good	110 (15.9)	48 (21.4)	2 (8.3)	34 (20.5)	19 (11.2)	2 (7.7)	3 (9.1)	1 (4.5)	1 (3.6)
Standard	347 (50.1)	133 (59.4)	15 (62.5)	87 (52.4)	87 (51.2)	12 (46.2)	4 (12.1)	6 (27.3)	3 (10.7)
Poor risk	234 (33.4)	41 (18.3)	7 (29.2)	45 (27.1)	64 (37.6)	12 (46.2)	26 (78.8)	15 (68.2)	24 (85.7)
Not assessable	2 (0.3)	2 (0.9)	0	0	0	0	0	0	0
NPM1 wt standard risk	180 (26)	59 (26.3)	10 (41.2)	50 (30.1)	43 (25.3)	8 (30.8)	4 (12.1)	6 (27.3)	3 (10.7)
Post-C1 response									
CR (excluding CRi)	538 (77.6)	202 (90.2)	23 (95.8)	157 (94.6)	156 (91.8)				
CRi	46 (6.6)	22 (9.8)	1 (4.2)	9 (5.4)	14 (14.4)				

NOTE: Data given as No. (%) unless otherwise indicated.
Abbreviations: C1, course 1; C2, course 2; CR, complete remission; CRi, complete remission with absolute neutrophil count < 1,000/μL or thrombocytopenia < 100,000/μL; MRD, measurable residual disease; RD, resistant disease (< 50% reduction in blast numbers with > 15% residual blasts); PR, partial response (≥ 50% reduction in blast numbers with 5% to 15% residual blasts); wt, wild type.

Table A3. Outcomes for Patients by Peripheral Count Recovery Response Combined With MRD Status

MRD status	No. (%CRi)	5-Year (3-year) OS	*P*	5-Year (3-year) CIR	*P*
All patients					
Post-C1					
CR *v* CRi	933/88 (9.4)	53 *v* 39 (60 *v* 46)	.002	50 *v* 43 (46 *v* 40)	.6
MRD- CR *v* MRD- CRi	407/39 (9.6)	63 *v* 63 (70 *v* 63)	.2	40 *v* 33 (35 *v* 33)	.7
MRD+ CR *v* MRD+ CRi	528/9 (9.3)	45 *v* 19 (52 *v* 33)	.001	58 *v* 53 (54 *v* 47)	.6
Post-C2					
CR *v* CRi	716/89 (12.4)	54 *v* 38 (59 *v* 46)	.02	51 *v* 47 (48 *v* 44)	.9
MRD- CR *v* MRD- CRi	449/54 (12.0)	63 *v* 52 (68 *v* 52)	.05	61 *v* 57 (59 *v* 57)	.9
MRD+ CR *v* MRD+ CRi	267/35 (13.1)	37 *v* 20 (46 *v* 40)	.3	45 *v* 40 (41 *v* 36)	.9
Standard risk NPM1 wt					
Post-C1					
CR *v* CRi	241/19 (7.9)	52 *v* 42 (64 *v* 56)	.16	58 *v* 66 (53 *v* 66)	.2
MRD- CR *v* MRD- CRi	100/11 (11.0)	60 *v* 64 (77 *v* 64)	.2	49 *v* 66 (41 *v* 66)	.07
MRD+ CR *v* MRD+ CRi	141/8 (5.7)	48 *v* 25 (55 *v* 50)	.4	65 *v* 69 (61 *v* 69)	.8
Post-C2					
CR *v* CRi	180/24 (13.3)	54 *v* 47 (63 *v* 47)	.3	58 *v* 43 (55 *v* 43)	.6
MRD- CR *v* MRD- CRi	118/16 (13.6)	63 *v* 61 (70 *v* 61)	.6	52 *v* 34 (48 *v* 34)	.5
MRD+ CR *v* MRD+ CRi	62/8 (12.9)	35 *v* 23 (50 *v* 23)	.10	70 *v* 67 (70 *v* 67)	.7

Abbreviations: C1, course 1; C2, course 2; CIR, cumulative incidence of relapse; CR, complete remission; CRi, complete remission with absolute neutrophil count $< 1{,}000/\mu L$ or thrombocytopenia $< 100{,}000/\mu L$; MRD, measurable residual disease; NPM1, nucleophosmin. OS, overall survival; wt, wild type.