Response assessment in acute myeloid leukemia by flow cytometry supersedes cytomorphology at time of aplasia, amends cases without molecular residual disease marker and serves as an independent prognostic marker at time of aplasia and post-induction

Thomas Köhnke,[1,2] Veit Bücklein,[1,2] Sandra Rechkemmer,[1] Stephanie Schneider,[1,3] Maja Rothenberg-Thurley,[1,4,5] Klaus H. Metzeler,[1,4,5] Maria-Cristina Sauerland,[6] Wolfgang Hiddemann,[1,4,5] Karsten Spiekermann[1,4,5] and Marion Subklewe[1,2,4,5]

[1]Laboratory for Leukemia Diagnostics, Department of Medicine III, University Hospital, LMU Munich, Munich; [2]Translational Cancer Immunology, Gene Center, LMU Munich, Munich; [3]Institute of Human Genetics, LMU Munich, Munich; [4]German Cancer Consortium (DKTK), Heidelberg; [5]German Cancer Research Center (DKFZ), Heidelberg and [6]Institute of Biostatistics and Clinical Research, University of Münster, Münster, Germany

Correspondence: MARION SUBKLEWE
marion.subklewe@med.uni-muenchen.de
doi:10.3324/haematol.2018.215236

SUPPLEMENTARY APPENDIX

Supplementary Results

Flow positivity during aplasia is an independent risk factor if patients with survival of less than 3 months are excluded

In order to delineate between early complications of therapy and prognosis due to residual disease, we excluded those cases where patients did not survive or were lost to follow-up within 3 months after initial diagnosis (Supplementary Figure S4). In these cases, we cannot exclude that the reason for death might be a complication of induction therapy not directly related to leukemia. For the remaining 125 cases, flow positivity during aplasia was associated with significantly shorter EFS (8.7 months vs. 19.1 months, p=0.004, Figure S5A) which was also true for the combined flow status (p=0.002, Figure S5B). In multivariate analysis, flow positivity during aplasia (HR 1.9, p=0.002, Table S4) and as a combined status (HR 1.9, p=0.001, Table S4) remained an independent risk factor next to age and ELN risk classification.

Supplementary Methods

Patient inclusion, diagnostic workup and treatment

From September 2012 to January 2016 patients with newly diagnosed AML (excluding M3) who received intensive induction chemotherapy and had comprehensive flow assessment available were included in our study. Patients provided written consent and collection of data was performed in compliance with the Declaration of Helsinki. Standard diagnostic work up at initial diagnosis included cytomorphology, cytogenics, mutation analysis and comprehensive phenotyping by flow cytometry. Results of cytogenetic and mutation analysis was stratified by Medical Research Council criteria[1] as well as by the 2017 European LeukemiaNet criteria[2]. All patients received intensive induction chemotherapy (either 7+3 or sHAM). Early response assessment during aplasia including detection of residual disease by flow cytometry was performed 7-9 days after completion of induction therapy.

Flow cytometric disease assessment

Bone marrow samples were analyzed by flow cytometry after purification of mononuclear cells (MNCs) by Ficoll density-gradient centrifugation. Leukemia-associated phenotypes were defined as previously described[3] using a comprehensive antigen panel (Supplementary Table S1). Flow analysis was performed using a

NAVIOS flow cytometer (Beckman Coulter, Brea, CA, USA). Data was analyzed with Kaluza software (Beckman Coulter, Brea, CA, USA).

Briefly, LAIPs were determined in primary diagnosis samples by gating on populations showing an aberrant antigen expression. The combination of gates applied to AML samples was also applied to bone marrow samples from healthy donors stained with the same antibody combinations. Only if less than 0.1% of MNCs from healthy donors displayed the putative LAIP phenotype, the population was defined as LAIP and the frequency as percentage of MNCs was determined in the leukemic sample (see Supplemental Figure S6 for frequency of LAIP phenotypes used in this study in healthy donor samples). As AML populations usually are heterogeneous, it is not possible to include the whole leukemic population into one LAIP. If possible, several LAIPs were defined per patient. In such cases, the LAIP with the highest value at each point of measurement was considered decisive and therefore used for analysis.

During follow-up, bone marrow samples were stained using the antibody combinations (Tube #1-#5, Table S1) showing the most informative phenotype(s) and ≥1e6 MNCs per staining combination were acquired to achieve maximum sensitivity. Due to limitations in cell yield during aplasia, we prioritized flow MRD testing over any genetic testing for this time point. In addition, samples with low cell yield after Ficoll were not processed further and thus not included in this study. Gates defined at primary diagnosis were applied to determine the frequency of MNCs expressing the LAIP. ≥0.1% of MNCs expressing the LAIP phenotype was defined as flow positive. All flow cytometry data was manually evaluated by experienced physicians trained in laboratory hematology to detect any sample artifacts and for the detection of any antigenic shift that might have occurred within the leukemic population. Within the cohort analyzed in this study, antigen shift could only be observed in one case. The gating strategy for 3 representative cases is outlined in Supplementary Figure S7.

Quality Control and Proficiency Testing

All assays were performed utilizing stringent quality control measures to ensure highly reproducible data generation for the entire duration of the study. Specifically, standard operating procedures (SOPs) outlining sample handling, instrument QC, reporting of results and deviations have been implemented for all diagnostic procedures including those reported in this study.

Our laboratory has been accredited by the DAkkS (Deutsche Akkreditierungsstelle GmbH) to perform diagnostic flow cytometry assays on blood and bone marrow specimens according to DIN EN ISO 15189:2007. This accreditation has been held for the entire duration of the study and was reconfirmed after routine inspections by external auditors. In addition, the implementation of a Quality Management System in the laboratory has been certified by the German certification body TÜV SÜD, specifically regarding “laboratory analyses including cytomorphology, flow cytometry and genetics for the diagnostics and therapy of hematological neoplasms and in clinical trials” since 2010. This certification has been reconfirmed multiple times following routine audits during the entire duration of the study.

Regarding instrument settings and quality control we have performed daily QC according to the manufacturer’s recommendations for the NAVIOS flow cytometer using Flow-Check Fluorospheres (Beckman Coulter, Brea, CA, USA) as well as weekly quality control using standardized blood samples (CD-Chex, Streck, La Vista, NE, USA). Diagnostic assays were only performed if the instrument passed these quality control measures. Flow cytometry plots for all measurements for all diagnostic assays including the flow MRD assays performed in this study were inspected and evaluated by a total of ≥3 individuals before the release of any diagnostic report: the technician acquiring the data and at least two physicians trained in laboratory hematology including the head of the flow cytometry unit or her deputy. The entire process of sample preparation, acquisition and evaluation was routinely evaluated by blinded interlaboratory proficiency testing. Our laboratory has been participating in the INSTAND (https://www.instand-ev.de/en.html) round robin testing for immunophenotyping (once per year), the UK NEQAS testing program for Minimal Residual Disease in B-ALL (4x per year), the UK NEQAS testing program for Minimal Residual Disease in AML (4x per year) and most recently our own interlaboratory testing program using a harmonized MRD panel within the AMLCG (Acute Myeloid Leukemia Cooperative Group) and SAL (Studienallianz Leukämie) study groups [4].

Molecular MRD assessment

Sample preparation and the conditions of the RT-PCR assay of *NPM1*mut were performed as previously described[5]. MRD levels of the samples were expressed as a ratio of the *NPM1*mut normalized to the housekeeping gene *ABL1* and divided by the *NPM1*mut/*ABL1* ratio of an internal calibrator (the OCI/AML3 cell line). Similarly,

sample preparation, primers and probes, and conditions of the relative quantitative RT-PCR assays of *CBFB-MYH11* and *RUNX1-RUNX1T1* were as previously described by Gabert et al.[6] and Schnittger et al.[7]. MRD levels of the samples were expressed as a ratio of the fusion transcript levels normalized to the levels of the housekeeping gene ABL1 and divided by the fusion transcript/ABL1 ratio of the reference cell line Kasumi-1. RT-PCR negative results were scored at a ratio of less than 0.0001 for *NPM1*mut, 0.001 for *CBFß-MYH11* and less than 0.0005 for *RUNX1-RUNX1T1*.

Statistics

Relapse-free survival (RFS) and Event-free survival (EFS) were calculated as described previously[2]; defined as the time from achievement of a remission until the date or relapse or death from any case for RFS and as the time the time of diagnosis to the date of primary refractory disease, relapse from CR/CRi or death from any cause for EFS. The Kaplan–Meier estimator and log-rank test were used to analyze survival data. Cox's proportional hazards regression model was used to determine the influence of individual factors in multivariate analyses. Pearson chi square test was used to compare categorical variables between MRD groups. All statistical analyses were performed using the IBM SPSS Statistics, Version 23 (IBM, Armonk, NY, USA).

Supplementary Table S1. Antibody Panel

Tube	**FITC**	**PE**	**ECD**	**PC7**	**APC**	**A750**	**PacBlue**	**KrO**
#1	CD34	CD64	CD14	CD33	CD2		CD65	CD45
#2	CD15		CD34	CD33	CD117	CD11b		CD45
#3	CD33	CD34		CD13	CD56		HLA-DR	CD45
#4	CD34	7.1	CD3	CD33	CD7	CD19		CD45
#5	CD38	CD135	CD34	CD33	CD123			CD45

Supplementary Table S2. Patient characteristics

	MRD analysis cohort (n=161)	flow MRD at Aplasia available (n=145)			flow MRD post Induction available (n=121)		
		MRD pos.	MRD neg.	*p*	MRD pos.	MRD neg.	*p*
Gender							
Male	85 (52.8%)	38 (58.5%)	39 (48.8%)	*0.244*	27 (49.1%)	33 (50.0%)	*0.921*
Female	76 (47.2%)	27 (41.5%)	41 (51.2%)		28 (50.9%)	33 (50.0%)	
Age at diagnosis							
Median (Range)	57y (20-82)	63y (20-82)	51y (21-76)		55y (20-82)	54y (21-81)	
< 60 years	88 (54.7%)	28 (43.1%)	52 (65.0%)	*0.008*	34 (61.8%)	43 (65.2%)	*0.704*
≥ 60 years	73 (45.3%)	37 (56.9%)	28 (35.0%)		21 (38.2%)	23 (34.8%)	
AML Type							
de novo	132 (82.0%)	51 (78.5%)	68 (85.0%)		47 (85.5%)	55 (83.3%)	
sAML	23 (14.3%)	11 (16.9%)	10 (12.5%)	*0.566*	7 (12.7%)	9 (13.6%)	*0.899*
tAML	6 (3.7%)	3 (4.6%)	2 (2.5%)		1 (1.8%)	2 (3.0%)	
Cytogenetics[a]							
favorable	16 (9.9%)	4 (6.3%)	12 (15.0%)		4 (7.3%)	12 (18.2%)	
intermediate	115 (71.4%)	44 (68.8%)	60 (75.0%)	*0.024*	43 (78.2%)	46 (69.7%)	*0.209*
adverse	29 (18.0%)	16 (25.0%)	8 (10.0%)		8 (14.5%)	8 (12.1%)	
2017 ELN risk stratification by genetics							
favorable	59 (36.6%)	14 (21.5%)	40 (50.0%)		20 (36.4%)	33 (50.0%)	
intermediate	55 (34.2%)	26 (40.0%)	25 (31.3%)	*0.001*	22 (40.0%)	22 (33.3%)	*0.305*
adverse	47 (29.2%)	25 (38.5%)	15 (18.8%)		13 (23.6%)	11 (16.7%)	
Induction therapy							
sHAM	126 (78.3%)	46 (70.8%)	67 (83.8%)	*0.061*	44 (80.0%)	55 (83.3%)	*0.295*
7+3	35 (21.7%)	19 (29.2%)	13 (16.3%)		11 (20.0%)	11 (16.7%)	
Induction result							
CR	104 (64.6%)	31 (47.7%)	61 (76.3%)		39 (70.9%)	58 (87.9%)	*0.020*
CRi	26 (16.1%)	10 (15.4%)	14 (17.5%)	*<0.001*	16 (29.1%)	8 (12.1%)	
refractory	31 (19.3%)	24 (36.9%)	5 (6.3%)		N/A		
Allogeneic SCT							
	72 (44.7%)	34 (52.3%)	33 (41.3%)	*0.184*	28 (50.9%)	28 (42.4%)	*0.351*
Follow-up (months)							
Median	13.7	11.7			13.7		

Abbreviations: sAML, secondary AML; tAML, treatment related AML; N/A, not applicable; ELN, European LeukemiaNet; CR, Complete remission; CRi, Complete remission with incomplete count recovery; SCT, Stem cell transplantation.

[a]Cytogenetics according to Medical Research Council criteria.

Supplementary Table S3. Multivariate Analysis for flow MRD status in patients w/o molecular MRD marker

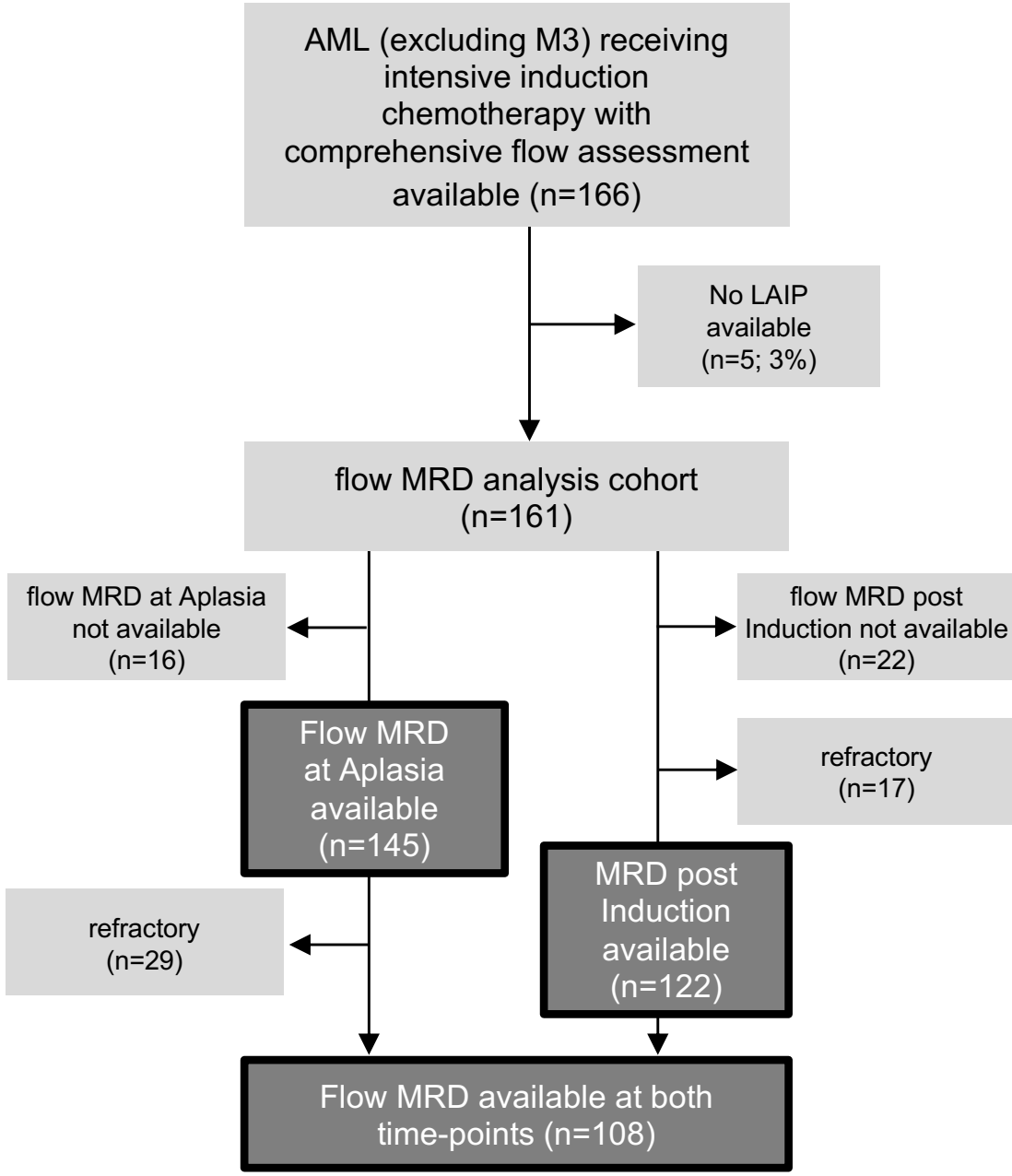
	HR (95% CI, p)	
	EFS	OS
flow MRD during aplasia and post induction available (n=54)		
2017 ELN risk stratification by genetics	p=0.8	p=0.6
favorable*	N/A	N/A
intermediate	1.4 (0.3-6.4, p=0.7)	0.4 (0.04-3.0, p=0.3)
adverse	1.6 (0.3-8.7, p=0.6)	0.4 (0.04-4.6, p=0.5)
Age (<60* vs. ≥60 years)	4.4 (1.7-11.7, p=0.003)	1.8 (0.3-10.4, p=0.5)
Combined MRD status (neg.* vs. pos.)	2.5 (1.2-5.4, p=0.016)	1.3 (0.4-4.3, p=0.7)

Abbreviations: CI, Confidence interval; ELN, European LeukemiaNet; EFS, Event Free Survival; HR, Hazard Ratio; MRD, Minimal Measurable Disease; OS, Overall Survival; *Reference category

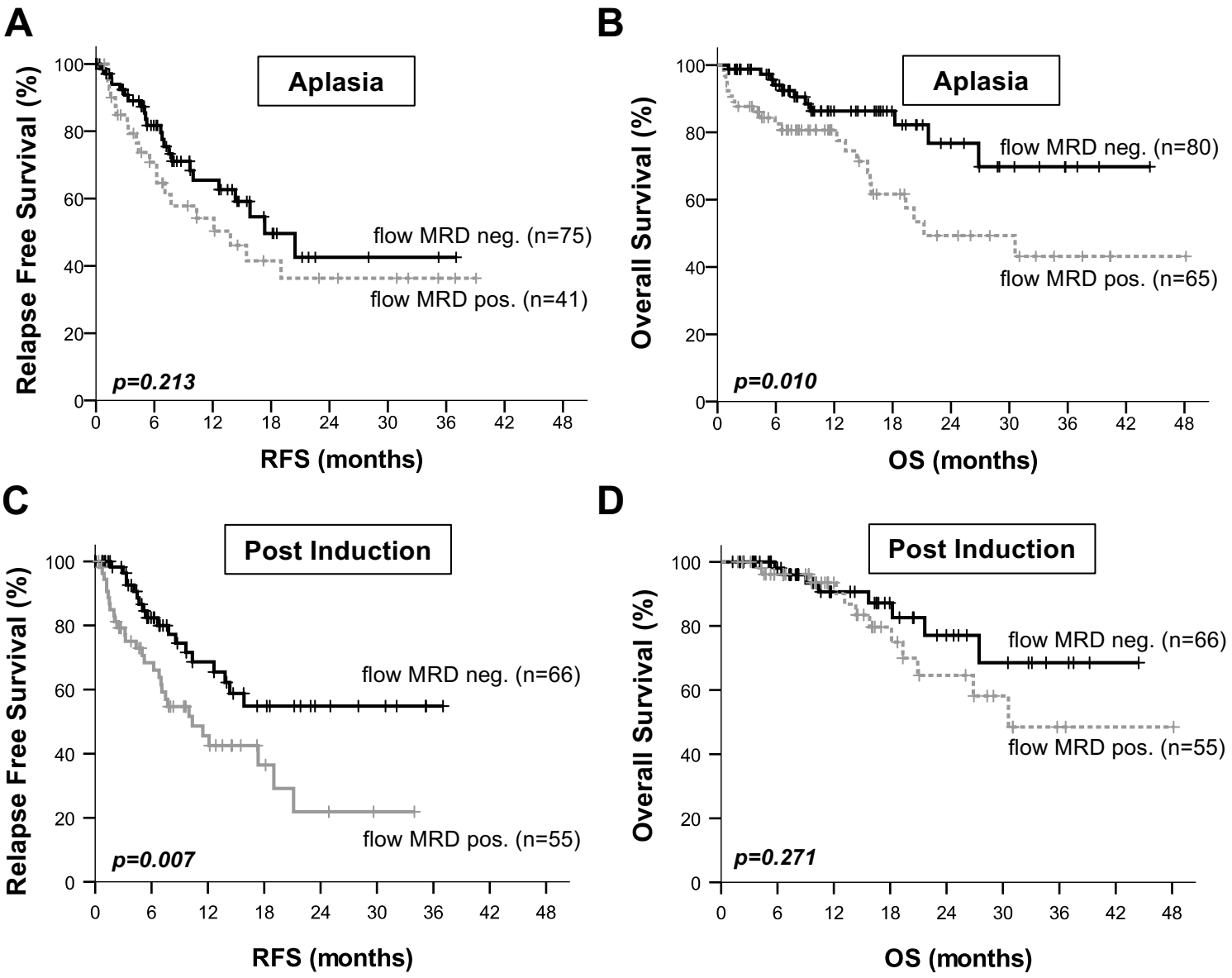
Supplementary Table S4. Multivariate Analysis for flow MRD status during aplasia if patients with survival of less than 3 months are excluded.

	HR (95% CI, p)	
	EFS	
flow MRD during aplasia available (n=125) or combined flow MRD available (n=125)		
2017 ELN risk stratification by genetics	p=0.002	p=0.348
favorable*	N/A	N/A
intermediate	1.1 (0.6-2.1, p=0.8)	1.3 (0.6-2.4, p=0.5)
adverse	2.7 (1.4-5.1, p=0.002)	1.8 (0.8-3.9, p=0.1)
Age (<60* vs. ≥60 years)	2.8 (1.7-4.8, p<0.001)	3.1 (1.7-5.7, p<0.001)
Flow status during aplasia (neg.* vs. pos.)	1.9 (1.7-3.2, p=0.020)	N/A
Combined MRD status (neg.* vs. pos.)	N/A	1.9 (1.3-2.8, p=0.001)

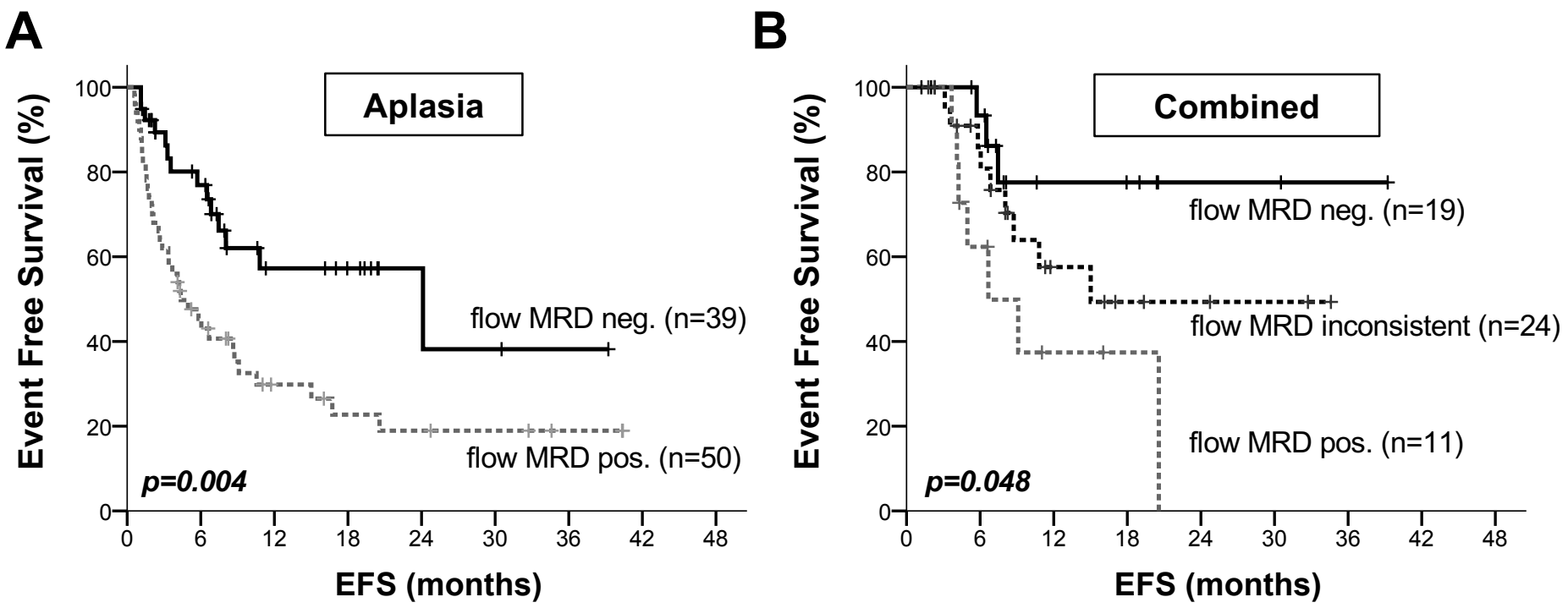
Abbreviations: CI, Confidence interval; ELN, European LeukemiaNet; EFS, Event Free Survival; HR, Hazard Ratio; MRD, Minimal Measurable Disease; OS, Overall Survival; *Reference category



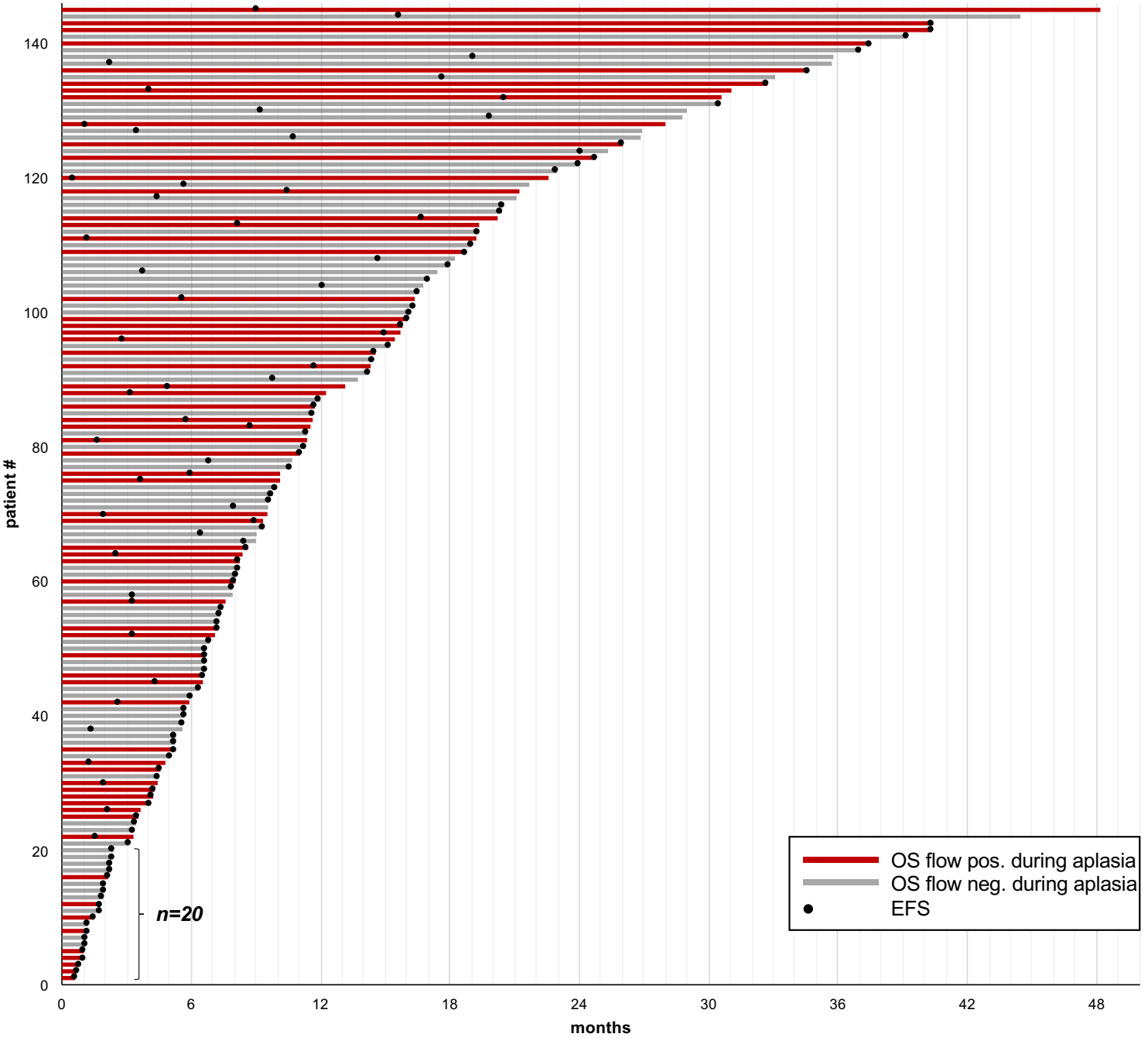
Supplementary Figure S1: Flow sheet depicting patient selection and data availability.



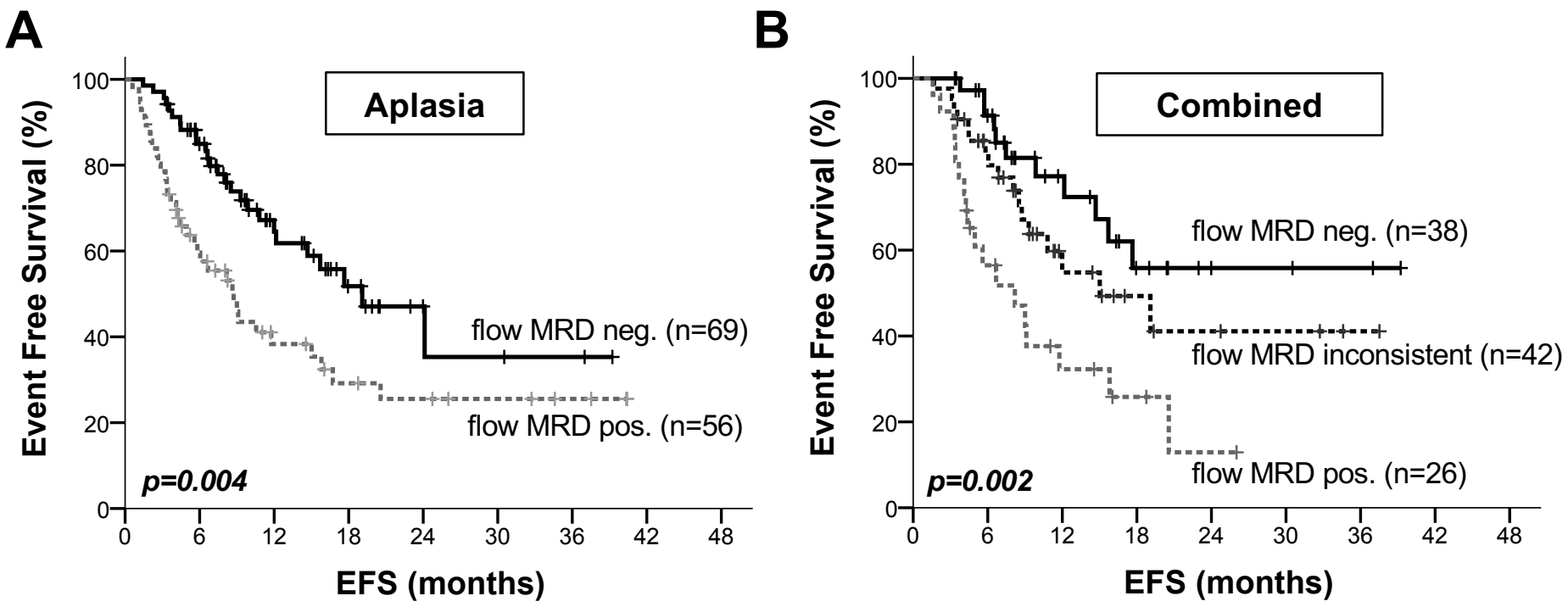
Supplementary Figure S2. Relapse Free and Overall Survival stratified by flow MRD. *A: Relapse free survival stratified by flow status during aplasia.* 41 patients were flow MRD positive and 75 patients were flow MRD negative. Median RFS was 13.8 months for flow MRD positive vs. 17.3 months for flow MRD negative patients (p=0.213). *B: Overall survival stratified by flow status during aplasia.* 65 patients were flow MRD positive and 80 patients were flow MRD negative. Median OS was 21.2 months for flow MRD positive vs. not reached for flow MRD negative patients. (p=0.010). *C: Relapse free survival stratified by flow MRD post Induction.* 55 patients were flow MRD positive and 66 patients were flow MRD negative. Median RFS was 10.3 months for flow MRD positive vs. not reached for flow MRD negative patients (p=0.007). *D: Overall survival stratified by flow MRD post induction.* Median OS was 30.6 months for flow MRD positive vs. not reached for flow MRD negative patients (p=0.271).



Supplementary Figure S3. Survival stratified by flow MRD for patients without a molecular MRD marker. *A: Event free survival stratified by flow status during aplasia in patients without a molecular marker (n=89).* 50 patients were flow MRD positive and 39 patients were flow MRD negative. Median EFS was 4.4 months for flow MRD positive vs. 24.1 months for flow MRD negative patients (p=0.004). *B: Event free survival stratified by combined flow MRD status in patients without a molecular MRD marker (n=54).* 11 patients were consistently flow MRD positive, 19 patients were flow MRD negative and 24 patients had inconsistent flow MRD measurements. Median EFS was 6.7 months for flow MRD positive, not reached for flow MRD negative and 15.0 months for flow MRD inconsistent patients (p=0.048).

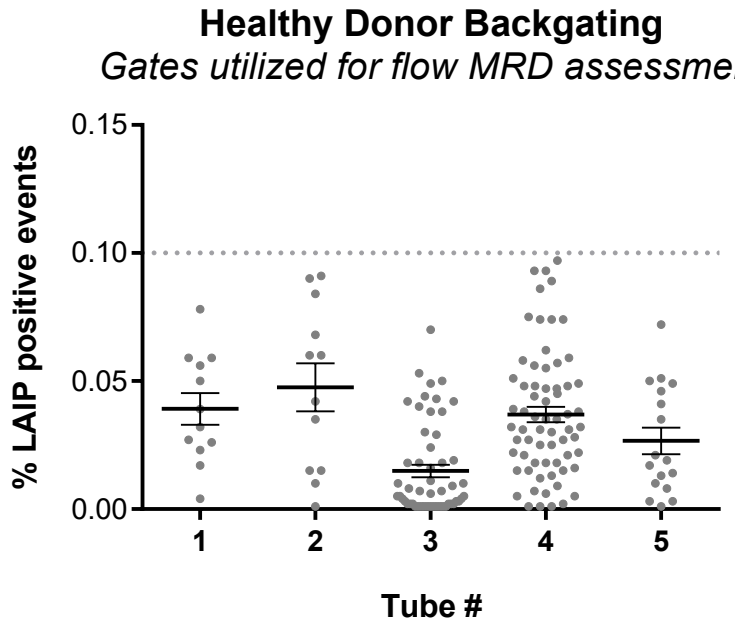


Supplementary Figure S4. Swimmer plots for patients with available flow status during aplasia.
Swimmer plot of OS (bars) and EFS (●) for patients who were either flow positive (red bars) or flow negative during aplasia (gray bars) during aplasia. The length of the bar corresponds to OS. Highlighted are 20 patients who either experienced an event or were lost to follow-up during the first 3 months after diagnosis.

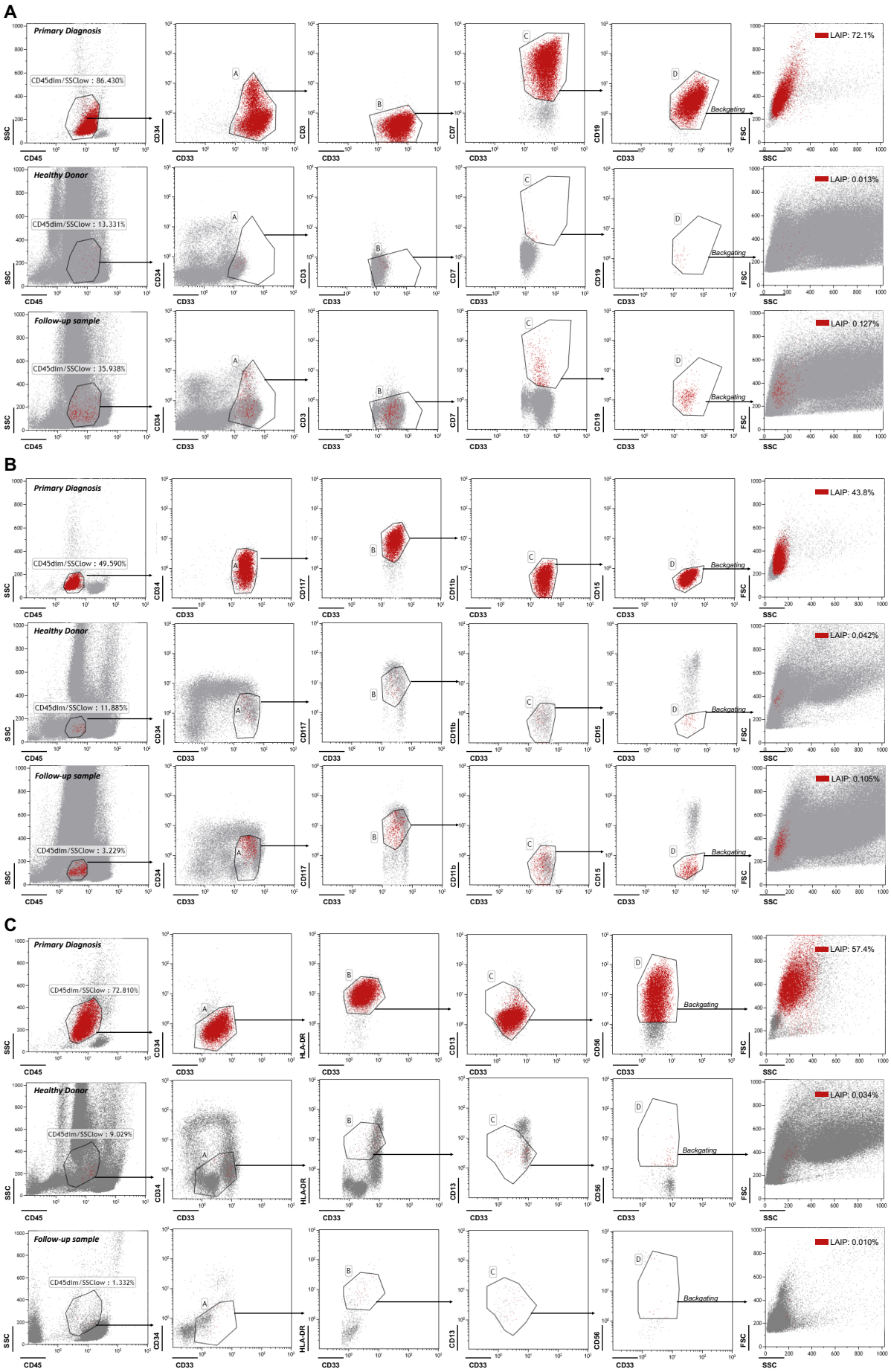


Supplementary Figure S5. Survival stratified by flow MRD status if patients with survival of less than 3 months are excluded.

A: Event free survival stratified by flow status during aplasia in our cohort if patients with survival of less than 3 months are excluded (n=125). 56 patients were flow MRD positive and 69 patients were flow MRD negative. Median EFS was 8.7 months for flow MRD positive vs. 19.1 months for flow MRD negative patients (p=0.004). *B: Event free survival stratified by combined flow MRD status in our cohort if patients with survival of less than 3 months are excluded (n=125). 26* patients were consistently flow MRD positive, 38 patients were flow MRD negative and 42 patients had inconsistent flow MRD measurements. Median EFS was 8.2 months for flow MRD positive, not reached for flow MRD negative and 15.0 months for flow MRD inconsistent patients (p=0.002).



Supplementary Figure S6. Frequency of LAIP phenotypes in healthy donor samples used in this study. During gating of patient samples (see Supplementary Methods for details), gating of putative LAIPs was adjusted to ensure that less than 0.1% of MNCs from healthy donors displayed the putative LAIP phenotype. Here, the frequency of LAIP positive events in healthy donor samples used in this study is presented.



Supplementary Figure S7. Gating strategy for 3 representative cases in this study. Gating was performed as described in Supplementary Methods by defining

the LAIP on primary diagnosis samples (first row in panels A, B, and C, respectively), ensuring that <0.1% of MNC in healthy donor bone marrow samples stained with the same antibody combination express the putative LAIP (middle row in panels A, B and C, respectively). Finally, during follow-up, flow MRD levels were determined by staining with the same antibody combination and applying the gates defined at primary diagnosis (bottom row in panels A, B and C, respectively).

Additional References

1. Grimwade D, Hills RK, Moorman AV, et al. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood*. 2010;116(3):354–365.
2. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424–447.
3. Köhnke T, Sauter D, Ringel K, et al. Early assessment of minimal residual disease in AML by flow cytometry during aplasia identifies patients at increased risk of relapse. *Leukemia*. 2015;29(2):377–386.
4. Buecklein V, Schnorfeil F, Koehnke T, et al. Flow cytometric assessment of minimal residual disease in AML: National harmonization in the HARMONIZE initiative. *Jahrestagung der Deutschen, Österreichischen und Schweizerischen Gesellschaften für Hämatologie und Medizinische Onkologie*. 2018;80.
5. Hubmann M, Köhnke T, Hoster E, et al. Molecular response assessment by quantitative real-time polymerase chain reaction after induction therapy in NPM1-mutated patients identifies those at high risk of relapse. *Haematologica*. 2014;99(8):1317–1325.
6. Gabert J, Beillard E, van der Velden VHJ, et al. Standardization and quality control studies of “real-time” quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia – A Europe Against Cancer Program. *Leukemia*. 2003;17(12):2318–2357.
7. Schnittger S, Weisser M, Schoch C, et al. New score predicting for prognosis in PML-RARA+, AML1-ETO+, or CBFBMYH11+ acute myeloid leukemia based on quantification of fusion transcripts. *Blood*. 2003;102(8):2746–2755.