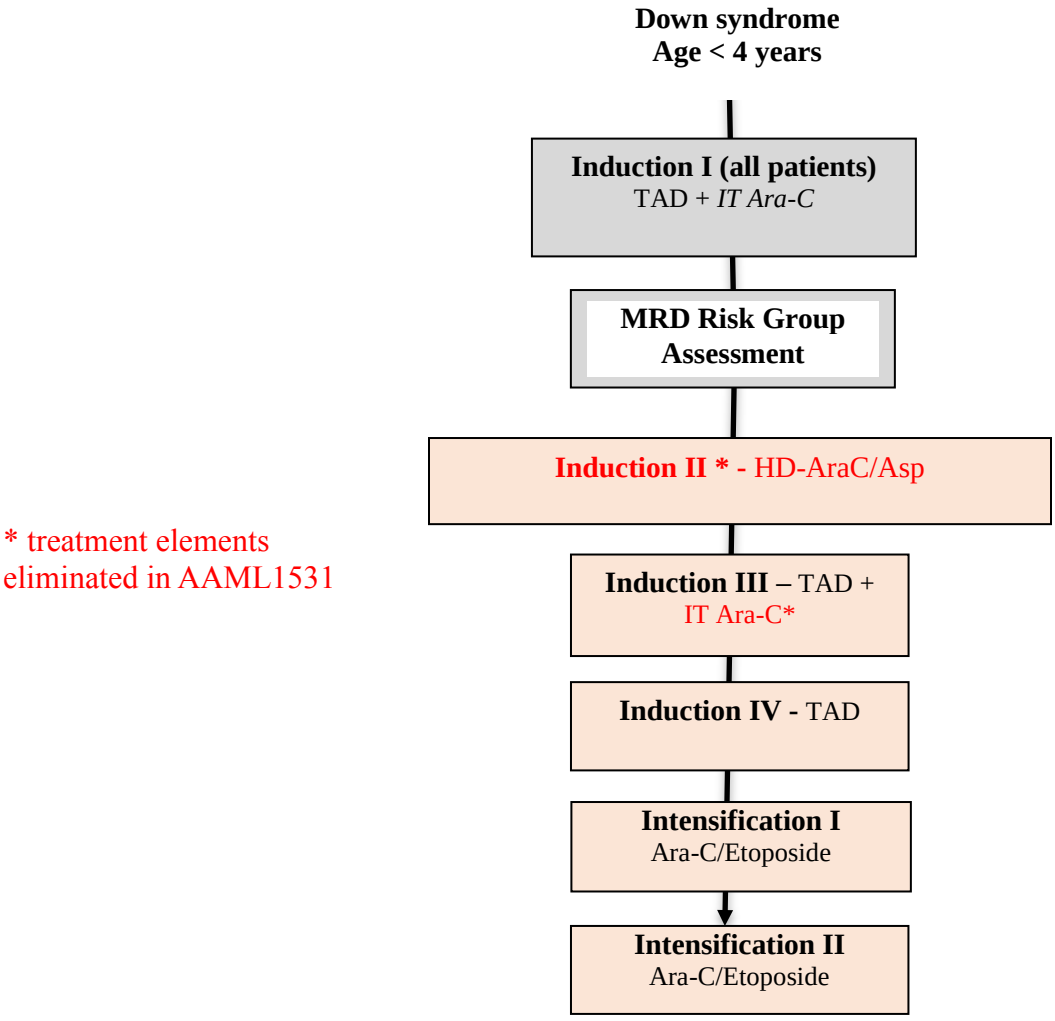
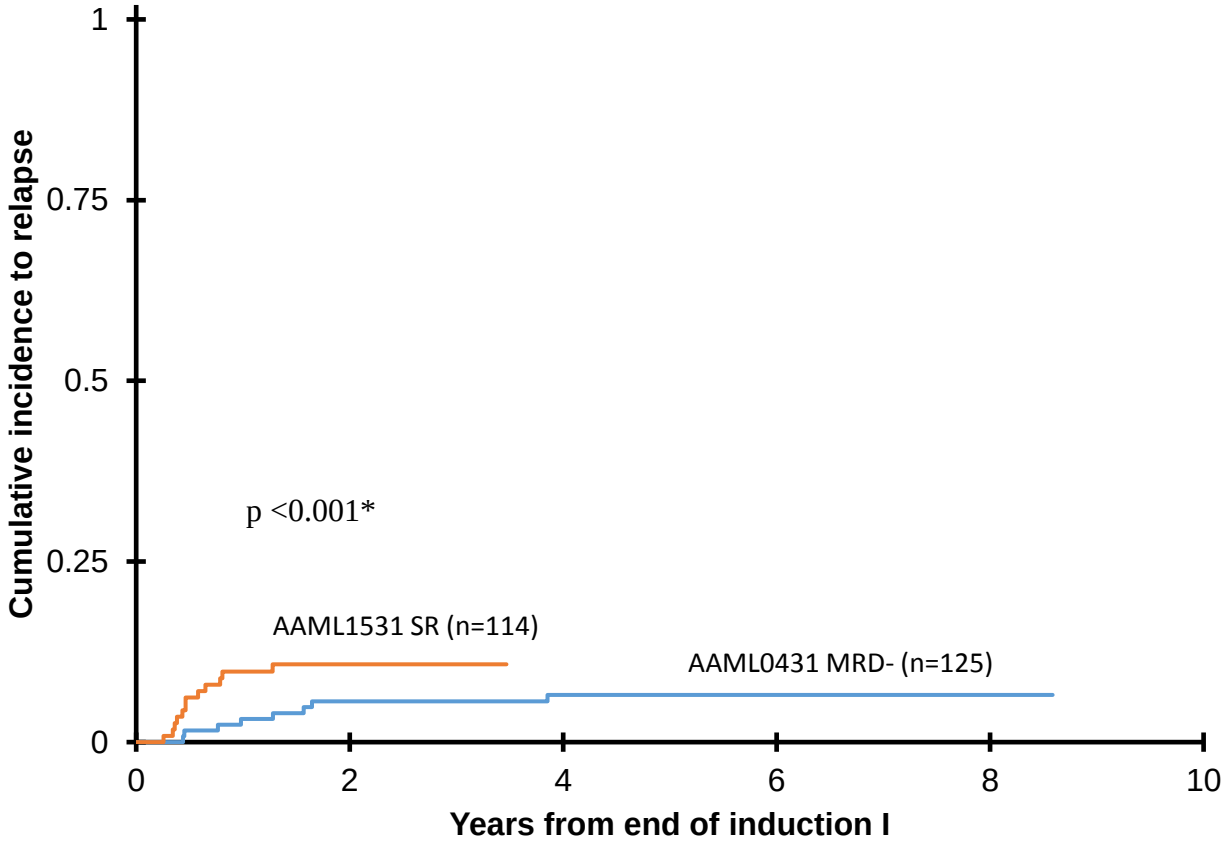
SUPPLEMENTAL DATA

Supplemental Figure 1 **Design of Predecessor Study AAML0431**



Post-Induction:
Induction II: Ara-C 3000 mg/m^2/dose every 12 hours on D1, 2, 8, 9 and *E.coli* asparaginase 6000 IU/m^2 D2 and 9
Induction III, IV: Ara-C 200 mg/m^2 (or 6.67 mg/kg) D1-4 by continuous infusion, Daunorubicin 20 mg/m^2 (or 0.67 mg/kg/d) D 1-4 (4 doses), Thioguanine 50mg/m^2 (or 1.65 mg/kg/dose) D1-4 (8 doses)
Intensifications I, II : Ara-C 100 mg/m^2/24h D1-7 continuous infusion, Etoposide 125 mg/m^2 D1-3 (3 doses)

Supplemental Figure 2 **Cumulative Incidence of Relapse**

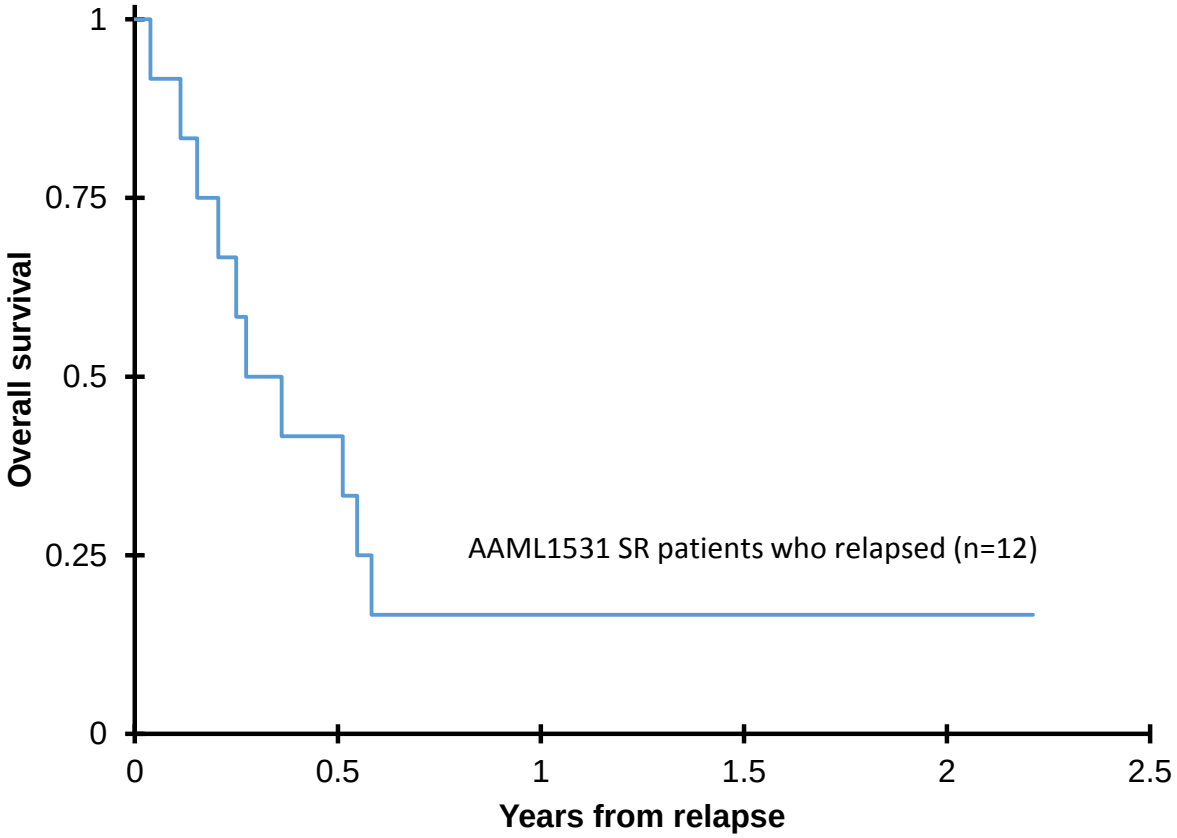


*Comparison at 2 years

	Number at risk at year								
Group	0	1	2	3	4	5	6	7	8
AAML1531 SR	114	100	52	9	0	0	0	0	0
AAML0431 MRD-negative	125	120	114	111	101	68	45	26	3

The cumulative incidence of relapse (95% CI) for AAML1531 SR patients was 10.8% (5.9% - 17.4%) at 2 years.

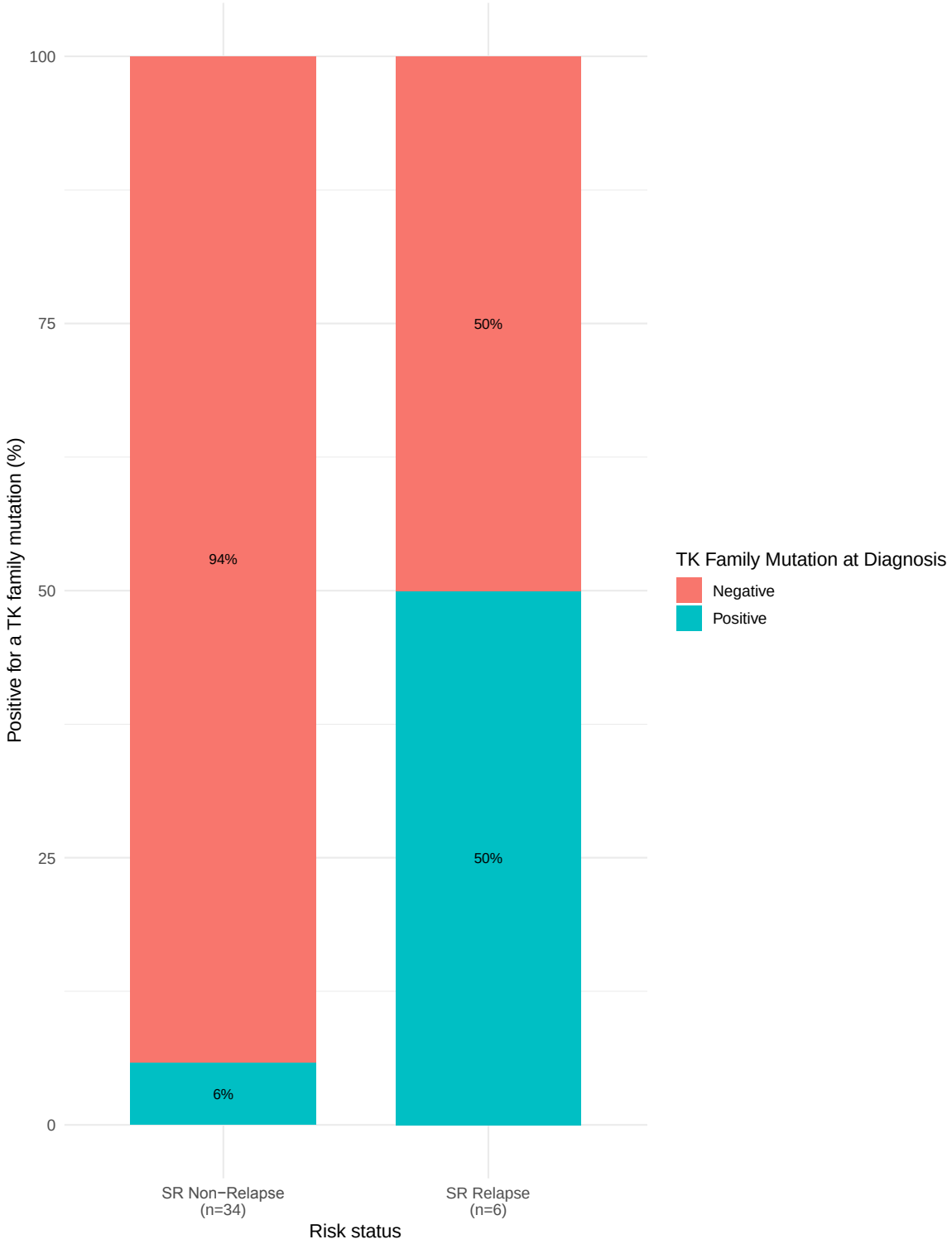
Supplemental Figure 3 Overall Survival of Standard Risk Patients after Relapse



	Number at risk at year					
Group	0	0.5	1	1.5	2	2.5
AAML1531 SR relapse patients	12	5	1	1	1	0

At 1 year the probability of survival (95% CI) was 16.7% (2.7%-41.3%) for AAML1531 SR patients who relapsed.

Supplemental Figure 4 **Mutations in tyrosine kinase family genes *JAK2, MPL, KIT* are more frequent in SR ML-DS patient who relapsed**



Mutations in tyrosine kinase family genes are found significantly more in SR patients that relapsed vs. those who did not (p=0.0178, two-sided Fisher's exact test).

Supplemental Table 1 Morphologic Diagnosis of SR ML-DS in AAML1531

Morphologic diagnosis[a]	**Acute myeloid leukemia[b]**	**Myelodysplastic syndrome**
Standard risk patients without relapse	40	35
Standard risk patients with relapse[c]	8	2
Total	48	37

[a] Of the 114 SR patients, 88 underwent pathology review, though 3 cases did not send adequate material for confirmation of diagnosis.
[b] Acute myeloid leukemia is defined as ≥20% blasts
[c] p-value 0.175

Supplemental Table 2 **Timing, Treatment and Outcome of Relapse Events in Patients with Standard Risk ML-DS**

Patient	Days since study entry	Treatment Phase	Relapse Site	Salvage Therapy	Alive Y/N
1	136	Intensification I	Bone Marrow	None	N
2	159	Intensification II	Bone Marrow	**HD-AraC**, anthracycline	N
3	169	Intensification II	CNS	None	N
4	189	Intensification II	Bone Marrow	None	N
5	203	Follow-up	Bone Marrow	**HD Ara-C**, mitoxantrone, gemtuzumab ozogamicin Refractory: decitabine, vorinostat SCT (no CR pre-SCT)	N
6	212	Follow-up	Bone Marrow	None	N
7	326	Follow-up	Bone Marrow	Chemotherapy (no HD-AraC, anthracycline, gemtuzumab ozogamicin)	N
8	327	Follow-up	Bone Marrow	**HD-AraC**, anthracycline SCT (MRD-neg. CR pre-SCT)	Y
9	197	Follow-up	Bone Marrow	**HD-AraC**, anthracycline	N
10	277	Follow-up	Bone Marrow	Chemotherapy (unknown if it contained HD-AraC, anthracycline, GO)	N
11	255	Follow-up	Bone Marrow	**HD-AraC**, anthracycline, gemtuzumab ozogamicin	N
12	511	Follow-up	Bone Marrow	**HD-AraC**, gemtuzumab ozogamicin	Y

Supplemental Table 3 Mutations in genes coding for signal transducers in ML-DS

Chr	Start	Ref	Alt	Gene	Type	AA change	COSMIC	VAF Replicate 1	VAF Replicate 2	Sample ID	Risk
chr9	5073770	G	T	JAK2	missense	p.V617F	COSM12600	0.287436	0.263685	PAXWJR	Relapse
chr9	5073770	G	T	JAK2	missense	p.V617F	COSM12600	0.095274	0.072727	PAYACG	Relapse
chr4	55575675	G	A	KIT	missense	p.A401T	NA	0.000843	0.00182	PAXWGU	Relapse
chr9	5073785	A	T	JAK2	missense	p.N622Y	NA	0.008395	0.014773	PAXLGI	No Relapse
chr1	43814979	G	A	MPL	missense	p.S505N	COSM27286	0.311336	0.287269	0BOK4C	No Relapse

Supplemental Table 4 Adverse Events by Course - SR Arm of AAML1531

	n	No. of patients with ANC recovery prior to starting next course	Days to ANC recovery, median (range)	ICU admission, % of patients	Febrile neutropenia ≥grade 3, % of patients	Sterile site bacterial infection ≥grade 3, % of patients
Induction I TAD	114	97	28 (0-47)	7.0	10.5	7.0
Induction II TAD	108	98	28 (0-42)	2.8	0.9	3.7
Induction III TAD	101	94	28 (0-38)	2.0	1.0	4.0
Intensification I araC/etoposide	98	88	31 (0-49)	1.0	3.1	5.1
Intensification II araC/etoposide	93	84	32 (0-71)	2.2	3.2	4.3

Supplemental Table 5: Comparison of Grade 3 or higher sterile site bacterial infections on AAML1531 and AAML0431

	AAML0431		AAML1531		
Chemotherapy Course	**Total N**	**Sterile Site Bacterial Infection grade >=3**	**Total N**	**Sterile Site Bacterial Infection grade >=3**	
	N	**N (%)**	**N**	**N (%)**	**p**
TAD #1*	204	39 (19.1)	191	12 (6.3)	**<0.001**
HD-AraC	199	45 (22.6)	n/a	n/a	
TAD #2	124	14 (11.3)	108	4 (3.7)	**0.031**
TAD #3	124	14 (11.3)	101	4 (4.0)	**0.044**
AE #1	123	17 (13.8)	98	5 (5.1)	**0.032**
AE #2	121	10 (8.3)	93	4 (4.3)	0.245

Induction II Block from AAML0431 (high-dose cytarabine) is omitted.
*TAD course 1 (Induction 1) includes all enrolled patients on both studies; subsequent courses include patients meeting AAML1531 definition for SR only.

Supplemental Table 6 Bacterial and Fungal Infections at Sterile Sites

Course		Organism	# of patients
Induction I	Bacteria	Streptococcus mitis	1
		Streptococcus pneumoniae	1
		Klebsiella pneumonia[&]	1
		Enterobacter cloacae[&]	1
		Staphylococcus epidermidis[*]	1
		Viridans group streptococcus	1
		Stenotrophomonas maltophilia	1
	Fungal	Mucor NOS[*]	1
		Trichophyton NOS[*]	1
Induction II	Bacteria	Stomatococcus mucilaginosus	1
		Bacillus NOS	1
		Streptococcus mitis[#]	1
		Granulicatella adiacens[#]	1
		Staphylococcus epidermidis[#]	1
	Fungal	Mucor NOS	1
Induction III	Bacteria	Staphylococcus epidermidis	1
		Enterobacter NOS	1
		Kocuria Rhizophila	1
Intensification I	Bacteria	Paenibacillus	1
		Actinomyces	1
		Abiotrophia defectiva	1
		Coagulase-negative staphylococcus	1
		Staphylococcus aureus	1
		Escherichia coli	1
	Fungal	Aspergillus terreus	1
Intensification II	Bacteria	Coagulase-negative staphylococcus	1
		Viridans group streptococcus	1

[&], [*] and [#] indicate isolates from the same patient

Supplemental Table 7 Clinically and Microbiologically Documented Viral Infections

Course		Virus	# of patients
Induction I	Respiratory	Coronavirus Human metapneumovirus Parainfluenza virus Rhinovirus/enterovirus	1 2 2 1
	Gastrointestinal	Norovirus Astrovirus	1 1
Induction II	Respiratory	Respiratory syncytial virus	1
Intensification I	Respiratory	Rhinovirus/enterovirus	1
Intensification II	Respiratory	Human metapneumovirus Rhinovirus/enterovirus	1 1

Supplemental Table 8 Cumulative Drug Doses

Doses mg/m^2	**AAML 1531** SR	**AAML 0431** (ref. 14)	**AML99 DS** (ref. 12)	**ML-DS 2006** (ref. 15)
Cytarabine	3,800	27,800	3,500	27,400
Daunorubicin	240	240		
Idarubicin				34
Mitoxantrone				14
Pirarubicin			250	
Total anthracycline*	240	240	250	226
Etoposide	750	750	2,250	450
Number of cycles i.t. doses	5 1	6 2	5 0	4 4

*using a conversion factor of 5 for Idarubicin, 4 for Mitoxantrone and 1 for pirarubicin

SUPPLEMENTAL METHODS

Trial Design. All patients received the same first course Induction I (TAD: thioguanine 50 mg/m^2/dose BID, day 1-4; cytarabine 200 mg/m2/24h continuous infusion, day 1-4; daunorubicin 20 mg/m^2 on day 1-4 over 1-15 min; and a single dose of age-based intrathecal cytarabine). For patients younger than 36 months dosing was based on weight (see appended protocol). After count recovery, but no later than 49 days from start of induction therapy, residual disease in the bone marrow was measured by multidimensional flow cytometry using a single reference lab (Hematologics, Inc., Seattle, WA). Patients with an MRD level below 0.05% were assigned to the Standard Risk arm (SR, Arm A). SR therapy consisted of two more courses of DAT (Induction II and III, without further intrathecal therapy) followed by two identical courses of intensification therapy (Intensification I and II: AE, cytarabine 100 mg/m^2/24h continuous infusion, day 1-7; etoposide 125 mg/m^2/day, day 1-3). Patients with an MRD level of 0.05% or greater were assigned to the High Risk arm (HR, Arm B). HR therapy consisted of Induction II (MA, mitoxantrone 0.4 mg/kg/day, day 3-6; cytarabine 33 mg/kg/dose every 12 hours, day 1-4), Intensification I (AE, cytarabine 33 mg/kg/dose iv every 12 hours, day 1-5; etoposide 5 mg/kg/day , day 1-5) and Intensification II (Capizzi II: cytarabine 100 mg/kg/dose every 12 hours, day 1, 2, 8, 9; asparaginase E. coli 200 IU/kg/dose or Erwinia 830 IU/kg /dose, day 2 and 9). Dosing by weight was used for patients younger than 36 months of age. Patients were not required to meet minimal organ function requirements.

Flow cytometry. Bone marrow samples were collected (in sodium heparin) at the end of the first course of induction therapy, submitted to a single central reference laboratory (Hematologics Inc., Seattle, WA) and stained with a standardized panel of monoclonal antibodies (CD45 [2D1 (BD), PerCP], CD34 [8G12 (BD), APC], CD11b [D12 (BD), PE], HLA-DR [L243 (BD), FITC], CD36 [FA6.152 (BD), FITC], CD38 [HB7 (BD), FITC], CD16 [3G8 (BD), FITC], CD13 [L138 (BD), PE], CD14 [MϕP9 (BD), FITC), CD33 [P67.6 (BD), PE], CD7 [4H9 (BD), FITC], CD56 [MY31 (BD), PE], CD117 [104D2 (BD), PE]), CD123 [9F5 (BD), PE], CD19 [4G7 (BD), FITC] designed to detect residual disease by using the difference from normal approach (ΔN)(24-29). ΔN flow cytometry uses a standardized panel of monoclonal antibodies, quantitative immunofluorescence, and multidimensional data analysis to define the composition of bone marrow specimens. This approach can identify all lineages and maturational stages of normal cells resident in a stressed bone marrow. The relationships and absolute amounts of each gene product identified by a select panel of monoclonal antibodies are invariant from individual to individual with assay stability in demonstrated in the reference laboratory over 15 years. Unlike the leukemia associated immunophenotype (LAIP) technique (43-46), ΔN does not require access to the diagnostic specimen and is not impacted by changes in phenotype that can occur following therapy. Specimens were processed as previously described (24, 30). A total of 200,000 events were collected and data analysis was performed with WinList software (Verity Software House) using Boolean gating to identify normal and aberrant cell populations. Detection of residual disease required a cluster of 40 cells with at least two phenotypic abnormalities > 0.5 decades ($\log_{10}$ units) disparate from the nearest normal counterpart, allowing for the detection of leukemia down to a level of ≥0.02% of total nucleated cells (24, 30). Our analysis of ML-DS samples paid particular attention to avoiding misclassification of non-leukemic myeloid progenitors co-expressing CD56 and CD34 as well as CD33 (in addition to

CD117, CD13, CD34, and CD45), which uniquely occur in regenerating normal bone marrow of children with DS after chemotherapy for AML or ALL (31, 32) as residual leukemic disease. All data were reviewed by two independent analysts, who were blinded to patient information and came to agreement on a patient's residual disease status prior to issuing a report. A level above the threshold of detection (0.02%) was chosen as clinical cutoff for report as a positive result (0.05% of total nucleated cells) to meet regulatory agencies requirements. Results were reported to the submitting institution through a web-based platform (Rave EDC, Imedidata). We also analyzed expression of antigens that may serve as targets for immunotherapy, such as CD33 (47) and CD123 (48), in diagnostic bone marrow specimens of patients with ML-DS . We compared mean fluorescence intensities between ML-DS blasts of 110 SR patients enrolled on AAML1531 and 110 randomly selected pediatric AML patients without DS (enrolled on COG study AAML1031; NCT01371981).

Error-Corrected Sequencing (ECS) of Patient-Specific *GATA1* mutations. Somatic mutations in the hematopoietic transcription factor *GATA1* are almost universally found in ML-DS, and therefore have the potential to serve as a molecular biomarker of MRD (7, 49). Variant calling at low levels (<0.01) by standard next-generation sequencing (NGS) is limited due to errors introduced by library preparation or the sequencer, thus reducing its utility in distinguishing rare MRD clones from noise (50). We therefore used error-corrected sequencing (ECS), which has a limit of detection of 0.0001 variant allele fraction (VAF), to quantify ML-blasts with *GATA1* mutations. Details of the ECS workflow have been comprehensively described in prior studies (51, 52). Following construction of error-corrected consensus sequences (ECCSs) from read families (≥3 reads), a position-specific error model was built on a binomial distribution to identify true clonal mutations. Briefly, 250 ng of genomic DNA was PCR amplified with Q5 High-Fidelity 2x Master Mix (New England BioLabs). Custom i5 and i7 adapters (Integrated DNA Technologies) were annealed onto the amplicons; custom i5 adapters contain a string of 16 random nucleotides (16N) which serve as unique molecular indexes (UMIs) with the following sequence: AATGATACGGCGACCACCGAGATCTACAC(N1:25252525)(N1)(N1)(N1)(N1)(N1)(N1)(N1)(N1)(N1)(N1)(N1)(N1)(N1)(N1)(N1)ACACTCTTTCCCTACACGACGCTCTTCCGATCT). The number of UMI-tagged libraries was quantified via the QX200 droplet digital PCR (ddPCR) platform with EvaGreen Supermix (BioRad). Following quantification, libraries were normalized to have 1M unique molecules per library thereby allowing for ~10x coverage for each error-corrected consensus sequence (ECCS). Libraries were sequenced on an Illumina MiSeq v2 instrument with the following settings: 2 x 144 paired-end, 8 courses Index 1, 16 courses Index 2. Approximately 5% of PhiX control DNA was spiked in for read diversity. Following sequencing, reads were demultiplexed using the i7 adapter containing 8N sample-specific sequences and then grouped by UMI into read families (no hamming distance was allowed to increase specificity). Deduplication and error correction was performed on read families (minimum three reads) to generate an ECCS. ECCSs were aligned to UCSC hg19/GRCh37 using Bowtie2 and processed with Mpileup using the parameters -BQ0 -d 10000000000000 to remove coverage thresholds. Each genomic position was modelled independently via binomial statistics to build a position and substitution-specific error model. Variants significantly above background for each position and substitution were then called as true positives.

SUPPLEMENTAL RESULTS

MRD by Error-Corrected Sequencing (ECS) of *GATA1* Mutations and Targeted Sequencing of ML-DS blasts. Targeting exons 2 and 3 of *GATA1*, we applied ECS to 18 paired bone marrow samples from SR patients (collected at diagnosis and EOI-1) to measure *GATA1* MRD ≥0.0001 VAF. *GATA1* mutations were detectable in 89% of paired diagnostic samples at an average VAF of 0.1924 (range: 0.0010 - 0.6158 VAF). *GATA1* MRD was detected in 33% of samples at EOI-1 at an average VAF of 0.0015 (range 0.00014 - 0.00664). The proportion of detectable *GATA1* MRD was 60% among SR patients who relapsed (n=5) vs. 23% who did not (n=13), a difference that was not statistically significant (p=0.2682, two-sided Fisher's exact test). The mean VAF of *GATA1* mutations between SR patients who did vs. did not relapse was neither significantly different at diagnosis (0.3003 vs. 0.1434 VAF; p=0.1852, two-sided Student's *t* test) nor at EOI-1 (0.0005 vs. 0.0012 VAF; p=0.4436, two-sided Student's *t* test). To investigate contributions to relapse, we used an ECS panel targeting 80 genes frequently mutated in adult and pediatric AML to interrogate diagnostic samples from 40 SR ML-DS patients, including 6 who relapsed. Genes encoding tyrosine kinases were more frequently mutated in SR patients who relapsed compared to those who did not (p=0.0178, two-sided Fisher's exact test). Specifically, two *JAK2* V617F mutations (0.0840 and 0.2756 VAF) and a *KIT* A401T mutation (0.0013 VAF) were detected at diagnosis among six SR patients who relapsed. In SR patients without relapse (n=34), a *JAK2* N622Y mutation (0.0116 VAF) and *KIT* S505N mutation (0.2993 VAF) were found (**Supplementary Figure 4 and Supplementary Table 3**).