Protocol Title: J1227 - Consolidation Therapy for Acute Myeloid Leukemia Guided by Leukemia Stem Cell Behavior

Principal Investigator: **B Douglas Smith, MD**

Investigators:
Mark Levis, MD, Ph.D.
Margaret M. Showel, MD
Richard J. Jones, MD
Amy Dezern, MD
Ivana Gojo, MD
Keith Pratz, MD
Gabriel Ghiaur, MD, Ph.D.
Douglas Gladstone, MD
Karen Mackey
Jackie Greer
Jackie Jung

Statistician: Gary Rosner, PhD

Table of Contents

Objective

1.1. Primary Objective

The primary objective of the trial is to compare the two-year relapse-free survival (RFS) of patients with acute myeloid leukemia (AML), presumed to be at high risk for relapse due to the presence of leukemia stem cells (LSCs) in their bone marrow at first complete remission (CR1), who receive either standard cytarabine-based chemotherapy or allogeneic stem cell transplantation (SCT).

1.2. Secondary Objectives

1. Compare the 2-year RFS survival of those without detectable LSCs in CR1 to those patients with detectable LSCs in CR1 who are randomized to cytarabine-based consolidation chemotherapy.
2. Describe the changes in the leukemic population in blood and marrow during and after treatment.

2. Background

2.1. Introduction

Although most patients with AML achieve CR following standard induction chemotherapy, the majority subsequently relapse and succumb to the disease.[1-3] Currently, cytogenetic and molecular aberrations are the best prognostic indicators for AML patients.[3-5] However, these factors predict primarily for groups of patients and cannot prognosticate accurately for individual patients. For example, core-binding factor (CBF) cytogenetic abnormalities are considered favorable, yet roughly half of these patients relapse.[6] Overall, between 40-60% of patients with favorable or intermediate-risk AML will experience relapse. Prognosis after relapse is poor with five-year overall survival rate of less than 10%.[7] A LSC paradigm may explain this failure of CR to reliably translate into cure. Leukemia appears to retain some semblance of the normal hematopoietic hierarchical structure: rare stem cells with self-renewal capacity give rise to partially differentiated progeny that comprise the bulk of the leukemia but possess only limited proliferative potential.[8] Although existing therapies are highly active against the leukemic bulk, it appears they spare the hardier LSCs responsible for relapse.[9;10] and OS and RFS of SCT in CR2 is much worse than in CR1 [11]

2.2. Leukemia Stem Cells

Since 1994 when Dick and colleagues reported that only rare AML cells, characterized by a classical $CD34^+CD38^-$ normal hematopoietic stem cell (HSC) phenotype, were capable of generating leukemia in immunodeficient mice,[12] these putative LSCs have been the focus of considerable research. Although it is generally accepted that $CD34^+CD38^-$ cells are enriched for LSCs,[12] this population is heterogeneous and includes both normal and leukemic cells. Moreover, recent data have challenged the $CD34^+CD38^-$ phenotype of LSCs in AML, leading many investigators to advocate for a functional definition of LSCs: those leukemic cells capable of engrafting immunodeficient mice.[13-16] However, even with this current gold standard, the identification of LSCs has proven elusive.[8;12] A significant portion of AML patient samples will not engraft mice, and the assay is cumbersome and often indeterminate.[17] More importantly, the clinical implications of this assay are unclear.[18] Furthermore, even in leukemia, where the cancer stem cell (CSC) model is perhaps best established, there is a paucity of data that LSCs are responsible for disease resistance or relapse.

Relapse after CR is due to minimal residual disease (MRD): residual leukemia cells persistent in quantities undetectable by conventional assays used to evaluate remission. If LSCs are indeed clinically relevant, then their persistence should correlate with recurrence, and MRD should be enriched for these cells. Exploiting the similarities between LSCs and their normal counterparts,[19] we recently employed strategies established for the isolation of normal HSCs[20-22] in chronic myeloid leukemia (CML). We found that the fraction of the $CD34^+CD38^-$ cells with high activity of Aldefluor, as determined by aldehyde dehydrogenase (ALDH) levels, was highly enriched for leukemic cells capable of engrafting immunodeficient mice.[23] Therefore, we assessed ALDH activity as a marker for clinically significant MRD in AML.

2.3. ALDH activity distinguishes LSCs from their normal counterparts

When analyzed for ALDH activity, the normal bone marrow $CD34^+CD38^-$ cells consistently exhibited two, non-overlapping populations: one expressing low ALDH activity ($CD34^+CD38^-ALDH^{low}$) and the second expressing high activity ($CD34^+CD38^-ALDH^{high}$) (Fig 1a). The normal marrow $CD34^+CD38^-ALDH^{high}$ cells comprised an average of 10% (range 9-12%) of the total $CD34^+$ cells and 76% (range 61-85%) of the $CD34^+CD38^-$ cells. As few as 1000 of these $CD34^+CD38^-ALDH^{high}$ cells engrafted NOG mice.

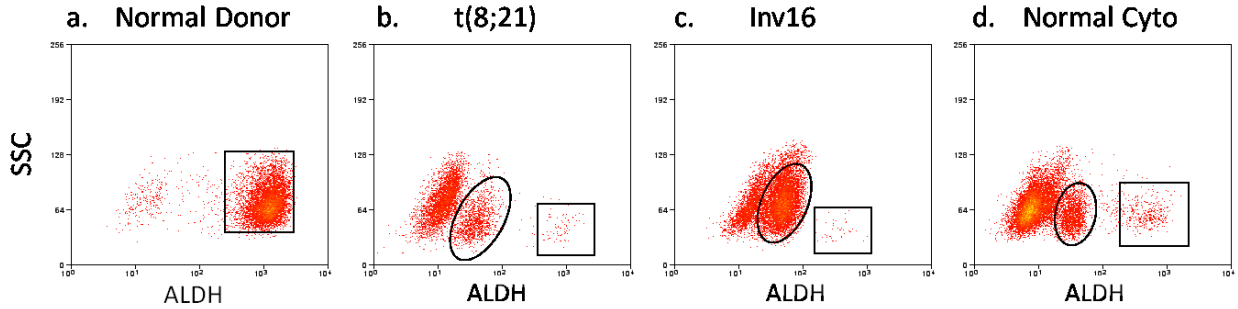

Figure 1. Expression of ALDH in the $CD34^+CD38^-$ population of bone marrow from a normal donor(a), and three leukemia patients with cytogenetic profiles of t(8;21), inv16, and 46xy profiles(b,c,d). Populations with high(boxed) and low ALDH expression are apparent in normal marrow(a). The AML samples contained a discrete third population of ALDH intermediate cells (circled). In the patients with CBF AML this $CD34^+CD38^-ALDH^{int}$ population was essentially completely leukemic by FISH, as was the $CD34^+CD38^-ALDH^{low}$ population. In contrast, the small $CD34^+CD38^-ALDH^{high}$ populations from CBF AML patients lacked the leukemia-specific FISH marker.

Initial AML analyses focused on patients with newly diagnosed CBF leukemias, as the FISH-detectable abnormalities enabled quantification of the percentage of leukemic cells in isolated populations. In contrast to the normal samples, the $CD34^+CD38^-$ cells from all CBF AML patients exhibited three well-defined populations by ALDH expression. They contained a population with intermediate ALDH expression ($CD34^+CD38^-ALDH^{int}$) (Fig 1b-c), in addition to high and low ALDH populations observed in normal marrow. The $CD34^+CD38^-ALDH^{high}$ cells were rare in the newly-diagnosed CBF AML patients, constituting an average of only 0.12% of the total $CD34^+$ cells (range 0.005-0.5%, $p<0.001$ vs. the normal samples) and 1.24% of the $CD34^+CD38^-$ cells (range 0.03-4.3%, $p<0.001$ vs. the normal samples). This $CD34^+CD38^-ALDH^{high}$ population was essentially devoid of cells with the leukemia-specific cytogenetic abnormality. Similar to those isolated from normal donors, as few as 1000 of these cells yielded normal human hematopoietic engraftment of NOG mice. Conversely, the intermediate and low ALDH fractions from were both virtually entirely leukemic by FISH. As few as 1000 of the $ALDH^{int}$ cells produced leukemic engraftment of NOG mice, whereas, the $ALDH^{low}$ cells did not engraft. The $CD34^+CD38^-$ cells from other cytogenetic variants of AML, including those with normal cytogenetics, also demonstrated three ALDH populations(Fig 1d) and similar FISH results. The $ALDH^{int}$ population appears to contain a putative LSC population, distinguishable from the normal HSC population.

This flow cytometric pattern containing an $ALDH^{int}$ population was absent only in four of 20 newly-diagnosed AML patients. In two of these patients, most of the $CD34^+CD38^-$ leukemic cells, as defined by FISH, exhibited high ALDH activity, with no discernible separate population of normal $CD34^+CD38^-ALDH^{high}$ cells. One patient had normal cytogenetics with a FLT3 internal tandem duplication, and the second had complex cytogenetics, including deletions of chromosomes 5 and 7. Notably, both patients had primary refractory disease, which ultimately proved fatal. In the two additional patients, the diagnostic leukemic cytogenetic marker was present only in the $CD34^-$ cells, as has been previously described in a minority of AML cases.[24] One of these patients had an 11q23 abnormality and has since relapsed within a year of diagnosis; the other has remained in CR for more than one year since diagnosis and nine months since allogeneic transplantation.

We hypothesized that, because this $ALDH^{int}$ population represents putative LSCs, if it is present during CR it would correlate with future relapse, and therefore define a clinically relevant MRD population.

Therefore, we evaluated AML patient samples at various stages of treatment to determine if their flow cytometric pattern predicted relapse.

2.4. Presence of LSCs following treatment predicts relapse

Of the 20 AML patients analyzed at diagnosis, three had primary refractory disease, one died during induction chemotherapy, one did not receive full induction or consolidation chemotherapy, and two others had $CD34^-$ leukemia. Of the 13 patients with $CD34^+$ leukemia analyzed at diagnosis who achieved morphologic CR after induction, follow-up samples were available in nine. An additional seven AML patients who achieved CR, but in whom diagnostic samples were not available, were also followed: two starting after induction and five after consolidation therapy, for a total of sixteen evaluable patients.
Of those 16 patients, 8 were analyzed in CR1 prior to consolidation. The CR samples exhibited two general patterns: two populations, with a predominant $CD34^+CD38^-ALDH^{high}$ population and a smaller $CD34^+CD38^-ALDH^{low}$ population, as seen in normal samples(Figure 2A); three populations including a $CD34^+CD38^-ALDH^{int}$ population (Figure 2B). Five patients exhibited the normal pattern, and both of the cell populations were normal by FISH. The three patients who have consistently exhibited this normal pattern remain in CR, with an average follow-up of 293 (range 185-370) days since diagnosis. In the other two patients the $ALDH^{int}$ population was detected at follow-up while still in CR after consolidation, and both ultimately relapsed. The remaining three patients exhibited an MRD pattern in their initial CR1 marrow (Figure 2B) and the $ALDH^{int}$ population was ≥85% leukemic by FISH. Two of these patients relapsed within 33 days of detection of the MRD pattern and subsequently died. The third patient underwent allogeneic SCT in CR1 due to adverse risk cytogenetics (complex karyotype, including deletion 7q) and has remained in CR for over 17 months.

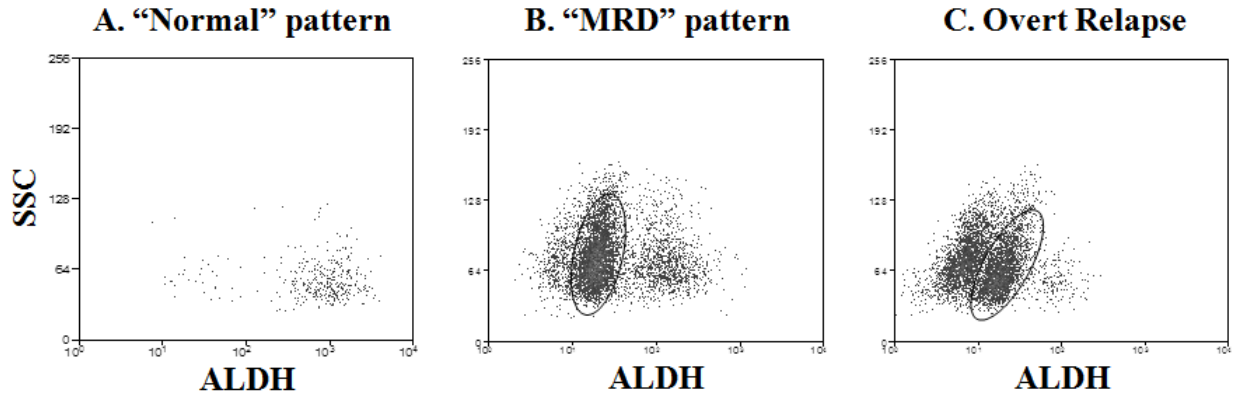

Figure 2. Aldeflour expression in the $CD34^+38^-$ bone marrow cells. Representative examples of staining patterns in CR1: (A) normal, (B) MRD, (C) overt clinical relapse.

Eleven AML patients who achieved morphologic CR with induction chemotherapy have been followed since completion of consolidation. Of the seven patients with a consistently normal flow pattern after consolidation, none has yet relapsed, with an average duration of follow-up of 509 days since diagnosis(range 185-810). However, all four of the patients with the MRD pattern after consolidation have relapsed at an average of 53 days (range 32-81) after detection of the MRD pattern ($p=0.003$ when compared to those patients with a consistently normal pattern). The $CD34^+CD38^-ALDH^{int}$ population was overwhelmingly leukemic (≥95%) by FISH. One of the relapsed patients converted to a normal pattern after re-induction and allogeneic transplantation and has remained in second CR for over one year. The MRD pattern persisted in two other patients, who achieved a second CR after re-induction, both of whom have subsequently relapsed. Of note, the $ALDH^{int}$ population was not detected in an additional two patients who have been in CR for more than three years, but have been followed only since their second year post-consolidation. The $CD34^+CD38^-ALDH^{int}$ population was present in all six analyzed cases of overt clinical relapse (Figure 2C).

2.5. Leukemic burden after chemotherapy is enriched for LSCs

According to the LSC paradigm, LSCs should constitute a small fraction of the leukemic burden during active disease, but should comprise the majority of the leukemic population during clinical remission. Although three patients in CR after induction who exhibited the MRD pattern had no morphologic, karyotypic, or FISH evidence of disease in the unfractionated marrow cells, the cytogenetic abnormality was still detectable in an average of 90% of the $CD34^+CD38^-ALDH^{int}$ cells (range 85-95%) and in an average of 22% of the $CD34^+$ cells (range 8-36%). This leukemic $ALDH^{int}$ population comprised an average of 34% of the total leukemic burden by FISH (range 9-51%). In contrast, this population constituted only 2% of the total leukemic burden at diagnosis (range 0.3-7%, $p<0.001$ vs. cytogenetic CR). Moreover, although the total leukemic burden decreased by 3-logs from diagnosis to CR ($p<0.001$), the $ALDH^{int}$ population decreased by only 1 log ($p=0.4$).

Three different patients, who had cytogenetic but no morphologic evidence of disease, had an average of 8% leukemic cells (range 4.5-9.5%) in their unfractionated marrow as compared to 96% in the $CD34^+CD38^-ALDH^{int}$ cells (range 94-99%, $p<0.001$). Furthermore, the proportion of the $ALDH^{int}$ fraction represented an average of 8% of the total leukemic clone (range 2-12%). At overt relapse, as at initial diagnosis, this population again comprised only a small fraction (average 1%, range 0.5-2%) of the total leukemic burden. These data demonstrated that this $CD34^+CD38^-ALDH^{int}$ population behaves as a putative LSC population.

Further, these data suggest that the presence of this $ALDH^{int}$ population predicts relapse, as expected of an LSC population. This would define a clinically relevant MRD population that indicates a high risk of AML relapse. Patients with this population, despite lacking the standard poor-risk prognostic markers, would be at a high risk of relapse and would likely benefit from more intensive therapy in first complete remission, due to the very poor prognosis of AML in relapse: five-year overall-survival rates are less than 10% after relapse.[7] Therefore, we propose a study of patients without poor-risk AML, as defined by current standards, who, upon achieving a CR after induction therapy, will receive consolidation treatment guided by the presence of LSCs in CR1.This will be a prospective, randomized, clinical trial, and patients who have the $CD34^+CD38^-ALDH^{int}$ population will be randomized to either standard chemotherapy *(Arm A)* or SCT *(Arm B)* for consolidation. Patients without LSCs in CR1 will receive standard consolidation chemotherapy *(Arm C)*.

3. Patient Selection and Enrollment

3.1. Inclusion Criteria

1. Age $\geq$ 18 years
2. Ability to give informed consent
3. New diagnosis of AML other than APL or poor-risk AML as defined in section 3.2

3.2. Inclusion Criteria Definitions

The original diagnosis of AML must have been confirmed by bone marrow aspirate and/or biopsy review by a JH Hematopathologist. Patients are eligible if they have AML that is not classified as poor-risk or APL. Poor-risk AML is defined as therapy-related, arising from a previous marrow disorder, or *de novo* AML associated with any of the following characteristics: trilineage dysplasia, FLT3/ITD mutation, poor-risk cytogenetics including: chromosome 3, 5, or 7 abnormalities, t(6;9), and complex karyotype.

3.3. Exclusion Criteria

1. Has already had a bone marrow biopsy and aspirate to assess remission status after induction therapy
2. Any debilitating medical or psychiatric illness that would preclude ability to give informed consent or receive optimal treatment and follow-up
3. Pregnancy: Women of childbearing potential who are β- HCG+

3.4. Inclusion of Women and Minorities

The proposed study is opened to both men and women, and to all racial/ethnic subgroups. There is no explicit mention of different treatment effects in male and female patients and in different racial/ethnic subgroups in the literature. Therefore, this study will not have separate accrual targets for these groups.

3.5. Donor Eligibility Criteria

Donor eligibility will be determined per standard BMT criteria.

3.6. Informed Consent

All patients eligible for the study must be evaluated by one of the study investigators.
Informed consent must be obtained and the consent form signed. A Johns Hopkins On-Study Form will be completed following fulfillment of the on-study requirements for laboratory work and eligibility criteria. For enrollment of non-English-speaking candidates, the JH IRB-mandated policies and procedures will be followed as listed at the following website:
https://irb.jhmi.edu/Guidelines/nonenglishconsent.html

4. Treatment Plan

4.1. Preconsolidation Plan

A bone marrow aspirate or biopsy confirming morphologic complete remission (< 5% blasts and normal cytogenetics and/or FISH) will be documented in the electronic medical records. Bone marrow aspirate / biopsy documenting complete remission and evaluations for consolidation registration (3.2) will be done within 2-4 weeks of the start of consolidation therapy. Patients with a bone marrow aspirate that is not evaluable for LSCs will be removed from the study.

4.2. Treatment Schedule

Patients without detectable LSCs by flow cytometry, will be assigned to standard cytarabine-based chemotherapy as per institutional standards. Patients with detectable LSCs by flow cytometry will be randomized to either standard cytarabine-based consolidation therapy or allogeneic SCT, per institutional standards. Patients will receive the transplant from a related donor who is at least a haploidentical match. The choice of the optimal donor will use institutional priority extant at the time. Patients eligible for myeloablative conditioning will be prioritized to receive it; patients not eligible for myeloablative conditioning will receive non-myeloablative conditioning.

4.3. Duration of Follow-up

Patients will be followed for up to five years from start of consolidation. Patients are routinely followed for by JH physicians for their post SCT or post-chemotherapy monitoring and patients who complete the study and remain with their care at JH will have survival captured at least annually.

4.4. Supportive Care Guidelines

Supportive care, including infection management and transfusion support, will follow good medical practice and institutional standard guidelines.

4.5. Study Parameters

There are institutional guidelines for study parameters for consolidation therapy with SCT or standard chemotherapy. Those guidelines will be followed per routine.

5. Treatment Evaluation

Response criteria and measurement of effect will be per **t**he NCI criteria for relapse in AML. Relapse following complete remission is defined as:

1. Peripheral Blood Counts: reappearance of blasts in the blood.
2. Bone Marrow Aspirate and Biopsy: Presence of > 5% blasts, not attributable to another cause (e.g., bone marrow regeneration) or dysplasia in greater than 10% of any lineage with cytopenias in one or more lineages or presence of a cytogenetic abnormality consistent with MDS or AML.
3. Development of extramedullary disease.

6. Laboratory Correlative Studies

In order to evaluate the putative LSC population, we will study cells from patient bone marrow aspirate. A minimum of 100,000 34^+ cells is required for complete analysis. The 34^+ cells will be isolated by Ficoll density centrifugation separation, followed by selection using anti-CD34 magnetic beads. Cells will be analyzed and sorted based on expression of CD34, CD38, and Aldeflour. FISH analysis for cytogenetic abnormalities will be performed on sorted cells.

7. Statistical Methods

7.1. Study Design and Objectives

This is designed as a prospective, randomized, clinical trial comparing two consolidation schema: conventional cytarabine-based consolidation and allogeneic blood or marrow transplantation (BMT). Patients will be in first complete remission (CR1) after cytarabine-based induction therapy. The randomized cohort of patients will be those with detectable leukemia stem cells (LSCs) in their remission marrow, whom we believe to be at high-risk for disease relapse.

7.1.1. Primary Objective

The **primary objective** of the trial is to compare the treatment-specific relapse-free survival (RFS) experiences of AML patients in CR1 with detectable LSCs, as described in the previous paragraph. The primary comparison will be with respect to two-year RFS. We chose two-year RFS, based on historical information that indicates that despite crossing hazards, the benefit of BMT to AML patients is realized by two years and extends beyond that point.[25;26] RFS is defined as the time until relapse/progression or death from any cause from randomization.

7.1.2. Primary Hypothesis

The primary hypothesis of the study is that allogeneic BMT will improve outcome (RFS) for AML patients in CR1 with detectable LSCs.

7.1.3. Secondary Hypothesis

A secondary hypothesis is that the presence of detectable LSCs will correlate with a poorer outcome for those AML patients randomized to conventional cytarabine-based consolidation therapy.

7.2. Sample Size and Power Considerations

The primary analysis will compare two-year RFS probabilities for the different groups. The comparisons are (1) BMT vs. non-BMT among LSC-positive patients and (2) non-BMT in LSC-positive patients vs. non-BMT in LSC-negative patients. The time point of 2 years was chosen based on data that indicate that most events with standard consolidation and allogeneic BMT occur by 2 years. The targeted sample size is 80 patients, 40 per treatment arm. The baseline RFS at 2 years is assumed to be approximately 10% for poor-risk AML patients receiving standard cytarabine-based consolidation. In actuality, the survival in our series of patients with detectable LSCs who received conventional-dose cytarabine-based consolidation was 0, but this was based on only a small number of patients. Furthermore, we assume that allogeneic BMT is associated with at least a 50% 2-year RFS in poor-risk AML patients in CR1. The assumed RFS distributions are shown in figure 3. The targeted sample size of 40 patients per group is sufficient to provide over 90% power to detect a difference in 2-year RFS (10% versus 40%), based on 5000 simulations from the assumed RFS curves, testing (at the 2-sided 0.1-level of significance) the equality of the Kaplan-Meier estimated treatment-specific 2-year RFS and associated standard errors. Even with the logrank test, we will have around 84% power to detect this difference. If, however, the 2-year RFS is 50% (the red curve in Fig 3), as we anticipate, we will have more than 90% power to detect this difference when testing the 2-year RFS or with a logrank test.

About 120 new AML patients are seen at Johns Hopkins per year. We anticipate about 1/3 (or 40 patients per year) will be poor-risk by standard cytogenetic/molecular criteria, so about 80 will be eligible for this study. If 3/4s agree to enroll on the study, 60 patients per year will be accrued. Our data suggest that half (30 per year, 15 per arm) will have detectable LSCs. Thus, the study will take about 3 years to accrue 40 patients to each arm. The 30 patients per year without detectable LSCs will serve as controls for the secondary objective. The final patient enrolled will be followed up for a minimum of 2 years, so that the targeted total study duration is 5 years (3 years accrual + 2 years follow-up on the last patient).

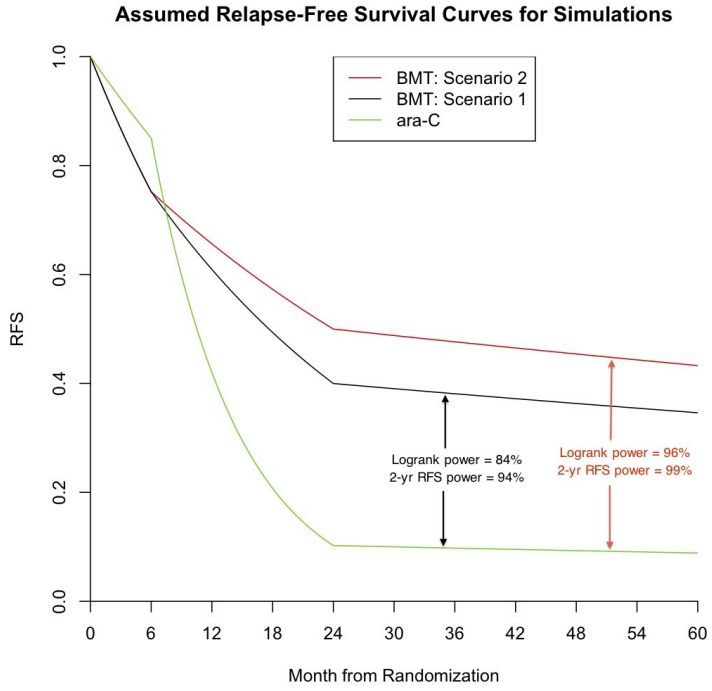

Figure 3. Assumed Relapse-Free Survival Curves

The study's secondary objective concerns comparing relapse-free survival among patients without detectable LSCs to the group of patients with detectable LSCs randomized to receive cytarabine-based consolidation . We anticipate roughly twice as many patients without detectable LSCs treated with cytarabine-based consolidation as with LSC, as shown in the previous paragraph. Therefore, we assumed that 40 LSC-positive patients will be randomized to receive cytarabine-based consolidation and that the study will enter 80 patients who are LSC-negative and will receive standard cytarabine-based consolidation. Assuming the same RFS curve for LSC-positive patients receiving cytarabine-based consolidation as in the sample size considerations for the primary objective (i.e., the randomized comparison), we will have roughly 85% power to detect a 55% reduction in the risk of relapse or death when comparing 40 LSC-positive patients getting cytarabine-based consolidation and 80 LSC-negative patients receiving cytarabine-based consolidation.

Table 1 shows various RFS values at 6-month increments up to two years from the start of treatment for various hypothetical RFS scenarios for the LSC-negative patients receiving cytarabine-based consolidation, relative to the hazard assumed for the LSC-positive patients randomized to receive cytarabine-based consolidation (i.e., the lowest curve in the figure 3). As with the primary objective, these calculations are based on 5000 simulations for each scenario with each test two-sided at the 10% level. In this case, the treatment-related piecewise exponential hazard functions are assumed proportional, unlike when comparing BMT to standard cytarabine-based consolidation. Nonetheless, we show power for both the logrank test (which is more powerful in this case) and a comparison of treatment-specific two-year RFS estimates.

Table 1. RFS Estimates

Relative Risk For LSC-Negative on cytarabine-based consolidation, Compared to LSC-Positive	RFS Probabilities After Start of Treatment				Power Compared to LSC-Positive	
	6 mo.	12 mo.	18 mo.	24 mo.	Logrank Test	Test 2-Year RFS
100%	0.85	0.42	0.21	0.10	0.10	0.10
75%	0.89	0.52	0.31	0.18	0.41	0.38
65%	0.90	0.57	0.36	0.23	0.66	0.60
60%	0.91	0.59	0.39	0.25	0.78	0.71
55%	0.91	0.62	0.42	0.29	0.88	0.83
50%	0.92	0.65	0.46	0.32	0.94	0.92

7.3. Interim Analysis and Stopping Guidelines

Interim analysis for efficacy and futility will be conducted annually by the SKCCC-appointed Data and Safety Monitoring Board (DSMB). The stopping guidelines serve as a trigger for consultation with the DSMB for additional review and are not formal "stopping rules" that would mandate automatic closure of study enrollment. Toxicity, adverse events, and other safety endpoints will be monitored regularly and reported to the DSMB at each interim analysis.

7.4. Secondary Objectives

1. The major secondary objective is to compare RFS of those without detectable LSCs in CR1 to those with detectable LSCs while in CR1 who are randomized to cytarabine-based consolidation chemotherapy. The analysis will compare RFS by the logrank test, as discussed at the end of section 7.2
2. The other secondary objective is to describe the changes in the clonogenic population, including changes in detectable cytogenetic abnormalities in blood and marrow, during consolidation (both cytarabine-based consolidation and allogeneic BMT)

7.5. Early Stopping Rules for Safety

This study will monitor acute grade III/IV GVHD in the transplant arm. If it becomes evident that the proportion of acute grade III/IV GVHD convincingly exceeds 10%, the study will be halted for a safety consultation. The stopping rule will hold enrollment if the posterior probability of toxicity risk exceeding 0.10 is 75% or higher. The prior for this monitoring rule is beta(1,9). This means that our prior guess at the proportion of acute grade III/IV GVHD is 10%, and there is 90% probability that this proportion is between 0.57% and 28.3%. The operating characteristics of the stopping rule are given in the Table 1 and are based on 5000 simulations:

Toxicity stopping rule:

- Serious AE in 2 out of 2 patients. Pr(Risk>0.1| Data) = 0.91
- Serious AE in 2 out of 3 patients. Pr(Risk>0.1| Data) = 0.889
- Serious AE in 2 out of 4 patients. Pr(Risk>0.1| Data) = 0.866
- Serious AE in 2 out of 5 patients. Pr(Risk>0.1| Data) = 0.842
- Serious AE in 2 out of 6 patients. Pr(Risk>0.1| Data) = 0.816
- Serious AE in 2 out of 7 patients. Pr(Risk>0.1| Data) = 0.789
- Serious AE in 2 out of 8 patients. Pr(Risk>0.1| Data) = 0.762
- Serious AE in 3 out of 9 patients. Pr(Risk>0.1| Data) = 0.902
- Serious AE in 3 out of 10 patients. Pr(Risk>0.1| Data) = 0.885
- Serious AE in 3 out of 11 patients. Pr(Risk>0.1| Data) = 0.867
- Serious AE in 3 out of 12 patients. Pr(Risk>0.1| Data) = 0.848
- Serious AE in 3 out of 13 patients. Pr(Risk>0.1| Data) = 0.828
- Serious AE in 3 out of 14 patients. Pr(Risk>0.1| Data) = 0.807
- Serious AE in 3 out of 15 patients. Pr(Risk>0.1| Data) = 0.786
- Serious AE in 3 out of 16 patients. Pr(Risk>0.1| Data) = 0.764

- Serious AE in 4 out of 17 patients. Pr(Risk>0.1| Data) = 0.888
- Serious AE in 4 out of 18 patients. Pr(Risk>0.1| Data) = 0.873
- Serious AE in 4 out of 19 patients. Pr(Risk>0.1| Data) = 0.858
- Serious AE in 4 out of 20 patients. Pr(Risk>0.1| Data) = 0.842
- Serious AE in 4 out of 21 patients. Pr(Risk>0.1| Data) = 0.825
- Serious AE in 4 out of 22 patients. Pr(Risk>0.1| Data) = 0.807
- Serious AE in 4 out of 23 patients. Pr(Risk>0.1| Data) = 0.789
- Serious AE in 4 out of 24 patients. Pr(Risk>0.1| Data) = 0.77
- Serious AE in 4 out of 25 patients. Pr(Risk>0.1| Data) = 0.75
- Serious AE in 5 out of 26 patients. Pr(Risk>0.1| Data) = 0.868
- Serious AE in 5 out of 27 patients. Pr(Risk>0.1| Data) = 0.855
- Serious AE in 5 out of 28 patients. Pr(Risk>0.1| Data) = 0.84
- Serious AE in 5 out of 29 patients. Pr(Risk>0.1| Data) = 0.825
- Serious AE in 5 out of 30 patients. Pr(Risk>0.1| Data) = 0.81
- Serious AE in 5 out of 31 patients. Pr(Risk>0.1| Data) = 0.794
- Serious AE in 5 out of 32 patients. Pr(Risk>0.1| Data) = 0.777
- Serious AE in 5 out of 33 patients. Pr(Risk>0.1| Data) = 0.76
- Serious AE in 6 out of 34 patients. Pr(Risk>0.1| Data) = 0.867
- Serious AE in 6 out of 35 patients. Pr(Risk>0.1| Data) = 0.854
- Serious AE in 6 out of 36 patients. Pr(Risk>0.1| Data) = 0.841
- Serious AE in 6 out of 37 patients. Pr(Risk>0.1| Data) = 0.828
- Serious AE in 6 out of 38 patients. Pr(Risk>0.1| Data) = 0.814
- Serious AE in 6 out of 39 patients. Pr(Risk>0.1| Data) = 0.8
- Serious AE in 6 out of 40 patients. Pr(Risk>0.1| Data) = 0.785

Table 2

True Tolerability Risk	Prob. Declare Treatment Not Tolerable	Avg. Sample Size
0.02	1.2%	39.6
0.05	9.4%	37.3
0.10	39.8%	29.9
0.15	72.9%	21.0
0.20	91.1%	14.4
0.25	98.0%	10.1
0.30	99.6%	7.7

Table 2. Operating characteristics of stopping rule based on 5000 simulations.

7.6. Early Stopping Guideline for Futility and Simulations for Each Cohort

The simulation studies demonstrate the power for the respective cohort sizes. Since we anticipate randomizing 30 patients per year and the primary endpoint of the study is 2-year RFS, 60 patients would be accrued by the time the first patient enrolled on study would reach two years of follow-up. Thus, almost all patients will have entered the study by the time we would have enough data to carry out an interim futility analysis. Therefore, the study does not include a formal futility monitoring rule. If, however, accrual is slower than we anticipate, we will consider including an analysis for futility.

8. Data Safety and Monitoring Plan

This is a DSMP Level I "Medium Risk" study under the SKCCC Data Safety Monitoring Plan (9/22/2011). The Clinical Research Office QA Group will perform an audit after the first subject has been treated and then periodically depending on the rate of accrual and prior audit results. All trial monitoring and reporting will be reviewed annually by the SKCCC Safety Monitoring Committee. The PI is responsible for monitoring the study. Data must be reviewed to assure the validity of data, as well as, the safety of the subjects. The PI will also monitor the progress of the trial, review safety reports, and clinical trial efficacy endpoints and to confirm that the safety outcomes favor continuation of the study.

8.1. Adverse Event Definition

An adverse event (AE) is defined as any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product, regardless of whether the occurrence is considered to have a causal relationship with treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, allogeneic SCT in this study, whether or not considered related to the medicinal (investigational) product. AEs are to be coded using an internationally recognized dictionary. This study will utilize the Common Toxicity Criteria (CTC) version 4.0 for toxicity where applicable for adverse event reporting. In cases where CTC cannot be applied to the toxic event, the investigator will quantify the toxicity based on intensity as defined below.

1. Mild: The subject is aware of signs or symptoms but they are easily tolerated; usually does not require additional therapy or discontinuation of investigational product.
2. Moderate: The signs and symptoms are sufficient to restrict but do not prevent usual activity; possible requires additional therapy but usually does not require discontinuation of investigational product
3. Severe: The Subject is unable to perform usual activities and usually requires discontinuation of investigational product.

8.2. Relationship to Investigational Product (SCT)

The investigator will classify the relationship of an adverse event to the investigational product according to the following definitions:

1. None: The time course between the investigational product and the occurrence or worsening of the adverse event rules out a causal relationship and or another cause is confirmed and no indication of involvement of the study product in the occurrence/worsening of the adverse event exists.
2. Unlikely: the time course between investigational product and occurrence or worsening of AE makes a causal relationship unlikely; the known effects of investigational product provide no indication of involvement in AE and another cause adequately explains the AE; regarding the AE, a plausible causal chain may be deduced from the known effects of investigational product but another cause is much more probably; or another cause is confirmed and involvement of investigational product in the AE is unlikely.
3. Possible: Regarding the AE, a plausible causal chain may be deduced, but another cause just as likely to be involved is also known; although there is no indication which possible cause is involved in the AE.
4. Probable: the course of the AE suggest involvement of the investigational product in the AE, although another cause cannot be ruled out.
5. Definite: the course of the AE indicates involvement of the investigational product in the AE and no indication of other causes exists.
6. Unclassifiable: only used for SAE: the available information is not sufficient for causality assessment.

8.3. Outcome

The investigator will record the outcome of the AE choosing one of the following categories

1. Recovered/resolved
2. Recovering/resolving
3. Not recovered/ not resolved

4. Recovered/resolved with residual effects as specified
5. Fatal
6. Unknown

8.4. Serious Adverse Event (SAE) Definition

As defined by the FDA CFR 312, a serious adverse event is one, occurring at any dose (including overdose), that results in any of the following:

1. Death
2. Life-threatening illness
3. Inpatient hospitalization or prolongation of existing hospitalization
4. Persistent or significant disability or incapacity
5. A congenital anomaly or birth defect
6. An important medical event
7. Pregnancy

8.5. Adverse Drug Reaction and Toxicity Monitoring

The study team (research nurse, study coordinator or attending) will assign toxicity scores using the NCI common Toxicity Criteria version 4.0, during consolidation therapy. A copy of the CTC version 4.0 is available at the CTEP home page (http://ctep.info,nih.gov). If an unexpected and serious toxicity occurs that would result in patients being subjected to unacceptable risk, the trial will be placed on hold while this toxicity is investigated.

8.6. Adverse Event Reporting

For study purposes, the following will be recorded and reported in accordance with IRB requirements:

- All unexpected serious adverse events

The PI will be responsible for reviewing the clinical course of all patients. All serious events will be confirmed as expected or unexpected by an expert panel of leukemia doctors. Reporting of serious adverse events will cease for subjects who are not following the protocol arm that they were assigned to.

In addition, the following will be tracked for study purposes and reported on a yearly basis to the IRB, or earlier if warranted:

- Any adverse events related to grades III-IV GVHD or to chronic GVHD in the first 2 years after BMT

All other adverse events will be neither tracked nor reported.

Reporting of subject deaths will cease in the following cases:

- Subjects on Arm A or Arm C who are greater than 30 days from the completion of consolidation chemotherapy
- Subjects on Arm B who are greater than 365 days from the date of bone marrow transplant
- Subjects who are not following the protocol arm that they were assigned to

8.7. Data Handling and Record Keeping

8.7.1. Case Report Forms

The PI and study coordinator will document in the patient files. Data required according to this protocol will be recorded on the CRFs developed by the PI. Entries on the CRF must be made with a ball-point pen and must be legible. Any documents related to the study must be archived at the study site or in a central archive. This includes the careful listing of the identities of the patients involved in the study. This list and the signed informed consent statements are key documents in the files to be stored by the PI. Patient

files will be archived according to local regulations. All documents related to the study must be retained until at least 15 years after the end of the study.

8.7.2. Patient Registry

The PI should maintain a patient registry of all patients entered into the study in the event that a safety issue arises after study completion.

9. Ethics

9.1. Institutional Review Board

The study protocol and any amendment that is not solely of an administrative nature must be approved by an Institutional Review Board (IRB).

9.2. Ethical Conduct of the Study

The study will be conducted in accordance with the ethical principles of the Declaration of Helsinki.

9.3. Evaluation of Benefits and Risks/Discomforts

9.3.1. Potential Benefits

Patients will receive evaluation and treatment of their malignancy as a result of participating in this trial. Moreover, our preliminary published data suggest that patients with detectable LSCs after induction therapy have a poor outcome with standard cytarabine-based consolidation therapy, and allogeneic SCT improves outcomes in poor-risk patients, as defined by standard cytogenetic/molecular criteria. Therefore in patients with detectable LSCs, SCT may delay or prevent relapse. Alternative approaches to entering this trial, standard of care cytarabine-based consolidation or allogeneic SCT, will also be discussed before the verbal and written consent.

9.3.2. Measures for Minimizing Risk

Side effects can be unpredictable in nature and severity, although all care will be taken to minimize them. If patients suffer any physical injury as a result of participating in this study, immediate medical treatment is available at the treatment center. Frequent blood work will be taken to monitor side effects. Although no compensation is available, any injury will be evaluated and treated in keeping with the benefits or care to which patients are entitled under applicable regulations.

9.3.3. Risks/Benefits Analysis

Data gathered from both clinical and laboratory evaluations in this trial will be analyzed frequently to ensure safety of patients. Any new or significant finding(s) found during the course of the research will be shared and explained to each participant since that may affect a patient's willingness to participate further. Patient's anonymity will be protected to the maximum extent in all publications and presentations that result from this research.

9.3.4. Patient Information and Consent

The investigator or consent designee will explain the nature of the study, its purpose and associated procedures, the expected duration, and the potential benefits and risks of participation to each patient prior to his/her entry into the study (i.e., before examinations and procedures associated with selection for the study are performed). Each patient will have ample opportunity to ask questions and will be informed about the right to withdraw from the study at any time without any disadvantage and without having to provide reasons for this decision. Following this informative discussion, a patient will be asked if he/she is willing to sign and personally date a statement of informed consent. Only if the patient voluntarily agrees to sign the informed consent statement and has done so, may he/she enter the study. The patient will receive a copy of the signed and dated informed consent form. The signed informed consent statement is to remain in the investigator's files. The informed consent form and any other written information provided to patients will be revised whenever important new information becomes available that may be relevant to the patient's consent, or there is an amendment to the protocol which necessitates a change to the content of the written informed consent form. The investigator will inform the patient of changes in a timely manner and will ask the patient to confirm continuation of his/her participation in the study by his/her signature on the revised informed consent form. Any revised written informed consent form must receive the IRB's approval/favorable opinion in advance of use.

9.4. Financial Disclosure

Each investigator (including the principal investigator and any sub-investigators) who is directly involved

in the treatment or evaluation of research subjects must disclose certain financial arrangements. There are no financial disclosures for this protocol.

10. Removal of Subjects from Study

Subjects will be taken off study for the following events:

1. Death
2. Completion of 5 years of follow-up
3. Diagnostic bone marrow aspirate, if available, has no LSCs. (Patients without an available diagnostic aspirate will stay on study.)
4. Recovery marrow after induction is not evaluable for LSCs
5. Subject is not in CR1 after induction
6. Prior to randomization, subject is found to have no donor available for bone marrow transplant
7. Prior to randomization, subject is unable to proceed with study for any other reason

11. References

Reference List

1. 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. 2003;21:4642-4649.

2. Jemal A, Clegg LX, Ward E et al. Annual report to the nation on the status of cancer, 1975-2001, with a special feature regarding survival. Cancer 2004;101:3-27.

3. Dohner H, Estey EH, Amadori S et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood 2010;115:453-474.

4. Grimwade D, Walker H, Oliver F et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. Blood 1998;92:2322-2333.

5. Byrd JC, Mrozek K, Dodge RK et al. Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). Blood 2002;100:4325-4336.

6. Estey E, Dohner H. Acute myeloid leukaemia. Lancet 2006;368:1894-1907.

7. Rowe JM, Li X, Cassileth PA et al. Very Poor Survival of Patients with AML Who Relapse after Achieving a First Complete Remission: The Eastern Cooperative Oncology Group Experience. ASH Annual Meeting Abstracts 2005;106:546.

8. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat.Med. 1997;3:730-737.

9. Jones RJ, Matsui WH, Smith BD. Cancer stem cells: are we missing the target? J.Natl.Cancer Inst. 2004;96:583-585.

10. Huff CA, Matsui W, Smith BD, Jones RJ. The paradox of response and survival in cancer therapeutics. Blood 2006;107:431-434.

11. Villela LM, Bolanos-Meade J. Blood and Bone Marrow Transplantation for Acute Myeloid Leukemia. Clinical Leukemia 2009;3:E11-E21.

12. Lapidot T, Sirard C, Vormoor J et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature 1994;367:645-648.

13. Taussig DC, Miraki-Moud F, njos-Afonso F et al. Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. Blood 2008;112:568-575.

14. Taussig DC, Vargaftig J, Miraki-Moud F et al. Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34(-) fraction. Blood 2010;115:1976-1984.

15. Sarry JE, Murphy K, Perry R et al. Human acute myelogenous leukemia stem cells are rare and heterogeneous when assayed in NOD/SCID/IL2Rgammac-deficient mice. J.Clin.Invest 2011;121:384-395.

16. Eppert K, Takenaka K, Lechman ER et al. Stem cell gene expression programs influence clinical outcome in human leukemia. Nat.Med. 2011;17:1086-1093.

17. Pearce DJ, Taussig D, Zibara K et al. AML engraftment in the NOD/SCID assay reflects the outcome of AML: implications for our understanding of the heterogeneity of AML. Blood 2006;107:1166-1173.

18. Rombouts WJ, Martens AC, Ploemacher RE. Identification of variables determining the engraftment potential of human acute myeloid leukemia in the immunodeficient NOD/SCID human chimera model. Leukemia 2000;14:889-897.

19. Wang JC, Dick JE. Cancer stem cells: lessons from leukemia. Trends Cell Biol. 2005;15:494-501.

20. Christ O, Lucke K, Imren S et al. Improved purification of hematopoietic stem cells based on their elevated aldehyde dehydrogenase activity. Haematologica 2007;92:1165-1172.

21. Pearce DJ, Taussig D, Simpson C et al. Characterization of cells with a high aldehyde dehydrogenase activity from cord blood and acute myeloid leukemia samples. Stem Cells 2005;23:752-760.

22. Pearce DJ, Bonnet D. The combined use of Hoechst efflux ability and aldehyde dehydrogenase activity to identify murine and human hematopoietic stem cells. Exp.Hematol. 2007;35:1437-1446.

23. Gerber JM, Qin L, Kowalski J et al. Characterization of chronic myeloid leukemia stem cells. Am.J.Hematol. 2011;86:31-37.

24. van der Pol MA, Feller N, Roseboom M et al. Assessment of the normal or leukemic nature of CD34+ cells in acute myeloid leukemia with low percentages of CD34 cells. Haematologica 2003;88:983-993.

25. Slovak ML, Kopecky KJ, Cassileth PA et al. Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group Study. Blood 2000;96:4075-4083.

26. Cassileth PA, Harrington DP, Appelbaum FR et al. Chemotherapy compared with autologous or allogeneic bone marrow transplantation in the management of acute myeloid leukemia in first remission. N.Engl.J Med. 1998;339:1649-1656.