DEVELOPMENT OF A PROGNOSTIC INDICATOR FOR CURATIVE HEMATOPOIETIC STEM CELL TRANSPLANT REQUIREMENTS IN ACUTE MYELOID LEUKEMIA PATIENTS

DEVELOPMENT OF A PROGNOSTIC INDICATOR FOR CURATIVE HEMATOPOIETIC STEM CELL TRANSPLANT REQUIREMENTS IN ACUTE MYELOID LEUKEMIA PATIENTS

By Shiva Murali, B.M.Sc. (Honours)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University

Descriptive Note

McMaster University MASTER OF SCIENCE (2023) Hamilton, Ontario (biochemistry and Biomedical Science)

TITLE:	Development of a prognostic indicator for curative hematopoietic stem cell transplant requirements in acute myeloid leukemia patients
AUTHOR:	Shiva Murali, B.M.Sc. (Honours)
SUPRVISOR:	Dr. Mick Bhatia
NUMBER OF PAGES:	x, 36

Abstract

Acute myeloid leukemia (AML) is a deadly cancer of the blood and bone marrow defined by the accumulation of immature and non-functional myeloid progenitor cells. While AML is associated with a high success of chemotherapy-induced remission, it is accompanied by high relapse rates with poor response to subsequent therapies. Therefore, relapsed AML patients only have a 10% probability of long-term survival. An effective postinduction therapy is allogeneic hematopoietic stem cell transplantation (HSCT). However, complications associated with HSCT can be more severe than the AML disease itself. To date, no robust methodology is available to prospectively identify and distinguish AML patients that are more likely to benefit from HSCT. Our group has shown that AML patients with high leukemic progenitor cell content (LPC+) have a significantly lower overall survival (OS) when compared to patients with lower LPC content (LPC-). The objective of this study was to determine whether the LPC assay can be used as a functional predictor of post-HSCT survival. We hypothesized that LPC content correlates to post-HSCT survival times. We performed LPC assays on over 100 primary AML patient samples, showing that HSCT significantly improved OS in both LPC+ and LPC- patients, but LPC+ patients benefited more strongly than LPC- patients. This provides an initial basis to suggest that HSCT can offset the negative prognostic impact associated with high LPC content. To understand the biology of LPCs, we employed the Infinium HumanMethylation450 BeadChip assay to determine whether there are any methylation patterns that distinguish LPC+ and LPC- patients. However, we were not able to discover any uniquely methylated regions that separate the two groups, suggesting for further studies with an increased patient cohort, or extending the analyses to the transcript level. Given the rarity of curative approaches to cancers, a prognostic measure that could determine whether any single patient will benefit from HSCT will have an immediate impact.

Acknowledgements

I would like to express my sincerest gratitude to Dr. Mick Bhatia. Thank you for accepting me into your program despite my initial response in declining as I had accepted an offer at another institution. If it was not for the second chance you gave me in reconsidering my original decision, I would not have had this remarkable opportunity to pursue cutting-edge research in your laboratory. Your invaluable feedback and input throughout the course of my degree has helped me grow as a scientific researcher. Your high standards for research motivated me to constantly learn and always produce the highest quality work possible. I truly did not know what it meant to be a part of a "high-impact" laboratory until I joined the Bhatia Program. Thank you for giving me the opportunity to flourish in this fast-paced, yet educational environment. Finally, thank you for your continued support outside of the laboratory by helping me throughout my medical school application process in writing recommendation letters and signing countless forms.

I would like to extend my appreciation to my committee members, Dr. Anargyros Xenocostas and Dr. Andrew McArthur. Thank you for your enthusiasm and scientific input at each of my committee meetings. After each meeting, I learned something new about my own project that I had not considered previously. Without your support, this project would not have gotten to its current stage. Dr. Samantha Wilson, thank you for teaching me the ins and outs of DNA methylation analysis using R studio. As someone who had no experience in coding, I now leave my graduate studies with a valuable life skill that will certainly benefit me in my future career.

To all the members of the Bhatia Program – both past and present – thank you for making me feel welcome from the very first day. To Allison, thank you for being a second PI. Your expertise and approachability has been instrumental in my success and growth as a graduate student. Cam and Sean – the mentorship provided by the both of you from start to finish is the reason I have been able to successfully get to where I am today. In addition to training me on several lab techniques. thank you for patiently watching my practice presentations during late evenings and providing detailed feedback to ensure that I am fully prepared for the real day. Amro, David, and Caylie – thank you for helping me get my mind off lab work during long days as I truly enjoyed our impromptu chess matches in the office. And again to Cam – thank you for the countless early morning badminton games, as it helped me stay active outside the lab.

To Guha, thank you for being a great roommate and friend. I will always remember our late-night conversations and gaming sessions that helped get me through my time in Hamilton.

Sonia – thank you for being my biggest support outside of home. From 2nd year of undergrad to my final year of graduate studies, thank you for your constant words of encouragement during my most stressful times. Thank you for your understanding, and constantly adapting and compromising your expectations to prioritize my success.

Finally, to my parents and sister, thank you for all the support and belief you placed in me from the very beginning. Each of you played a role in my success behind the scenes – whether it was dropping me back and forth from Hamilton to Mississauga, making me food to save my time and ensuring I eat healthy, or having someone to constantly tease when I came home for the weekend. Thank you for always making me believe that my hard work will pay off as it truly has today.

Table of Contents

Descriptive Note ii
Abstract iii
Acknowledgements iv
List of Figures vii
List of Supplementary Figures vii
List of Tables viii
List of Supplementary Tables viii
List of Abbreviations ix
Declaration of Academic Achievement x
1.0 Introduction 1
1.1 Hematopoiesis 1
1.1.1 Hematopoietic Stem Cells 2
1.2 Acute Myeloid Leukemia 2
1.2.1 AML Prognosis 3
1.2.2. Heterogeneity in AML 4
1.2.3 AML Treatment 4
1.3 Hematopoietic Stem Cell Transplantation 5
1.4 Stemness in AML 7
1.5 DNA Methylation 9
1.6 Study Rationale 10
1.7 Objective and Hypothesis 10
2.0 Materials and Methods 11
2.1 Primary human AML patient samples 11
2.2 Leukemic Progenitor Cell Assay 12
2.2.1 Overview of the LPC Assay 12
2.2.2 Reliability and Validity of the LPC Assay 13
2.3 DNA Extraction 14
2.4 DNA Methylation Analysis 14
2.4.1 Illumina Methylation BeadChips Assay 14
2.4.2 Processing and analysis of methylation raw data 16
2.5 Statistical Analyses 17
3.0 Results 17

3.1 LPCs serves as a prognostic predictor of AML patient survival 18
3.1.1 Leukemic progenitor content emerges as a robust predictor of AML patient outcome 18
3.1.2 Leukemic progenitor content represented an independent prognostic factor when compared to clinical guidelines of AML patient risk stratification 21
3.2 Leukemic progenitor content correlates with post-HSCT survival 22
3.2.1 LPC+ vs LPC– Survival Outcomes Post-HSCT 22
3.3 No identifiable DNA methylation loci distinguishing LPC+ and LPC- patients 23
4.0 Discussion and Conclusion 24
Bibliography 29
Supplementary Data 33
Supplementary Figures 33
Supplementary Tables 34

List of Figures

Figure 3.1: Substantial variation exists between different functional and molecular measures of leukemic stemness 18

Figure 3.2: AML patient survival based on functional LSC and LPC content 20

Figure 3.3: Functional LPC frequencies represent an independent prognostic factor in human AML 22

Figure 3.4: LPC content measured by LPC assay predicts transplantation benefit in AML patients. 23

List of Supplementary Figures

Supplementary Figure 1: Overview of the LPC Assay 32

Supplementary Figure 2: Miniaturization and custom image analysis scripts allow high-throughput CFU testing. 32

List of Tables

Table 1. Clinical characteristics of 121 patients tested for functional LSC and leukemic progenitor content 12

Table 2. Clinical characteristics of 67 patients used for DNA methylation analyses 14-16

List of Supplementary Tables

Supplementary Table 1. Clinical annotation of patients used for HSCT survival studies 34-36

List of Abbreviations

A	AML	Acute myeloid leukemia
B	BH	Benjamini and Hochberg
	BM	Bone marrow
C	CFU	Colony forming unit
	CR	Complete remission
	CSC	Cancer stem cell
D	DNMT	DNA methyltransferase
E	EFS	Event-free survival
	ELN	European Leukemia Net
G	GVHD	Graft-versus-host disease
H	HSCT	Hematopoietic stem cell transplantation
	HSC	Hematopoietic stem cell
I	IDAT	Intensity data
L	LHSC	London Health Sciences Centre
	LPC	Leukemic progenitor cell
	LSC	Leukemic stem cell
M	MDS	Myelodysplastic syndrome
O	OS	Overall survival
P	PB	Peripheral blood
T	TBI	Total body irradiation

Declaration of Academic Achievement

This thesis was completed mainly by the work of Shiva Murali with the following contributions from the members of Dr. Mick Bhatia's Laboratory:

- Dr. Allison Boyd: provided scientific input, aided in data analyses, and pioneered the idea for using leukemic progenitor cell content as a prognostic tool for AML patient survival
- Justin Lu: Performed LPC assay experiments on numerous AML patient samples prior to my arrival in the Bhatia Program

1.0 Introduction

1.1 Hematopoiesis

Hematopoiesis is the process by which hematopoietic stem cells (HSCs) proliferate and differentiate to form the specialized adult blood cells[1]. The hematopoietic system is organized as a hierarchy, where HSCs reside at the top. Through a series of differentiation steps, HSCs eventually give rise to mature blood cells which include neutrophils, eosinophils, basophils, monocytes, lymphocytes, platelets, and erythrocytes[2]. These cells play crucial roles in maintaining life functions. Leukocytes contribute to the body's immune functions, erythrocytes are involved in oxygen transport, and platelets are responsible for wound healing. Located between the HSCs and mature blood cells in the hematopoietic hierarchy are progenitor cells. HSCs produce lymphoid and myeloid progenitor cells, which are multipotent in nature and have decreased self-renewal capacity[3]. Progenitor cells possess highly proliferative and developmental characteristics which allow for the rapid amplification and replenishing of the mature blood cells. Furthermore, these cells are also more abundant and more easily obtained than HSCs, giving them important therapeutic potential. Generally, as stem cells differentiate into progenitors, self-renewal capacity decreases and proliferative capacity increases[3]. This trend is not definitive as it is variable within the hierarchy. Progenitors that give rise to granulocytes—which have a high turnover rate—have a very high proliferative index. Whereas the common lymphoid progenitor—cell that gives rise to longer-lived B and T cells—has a comparatively lower proliferative index. This regulation of the hierarchy is important in maintaining homeostasis of the mature blood cells.

1.1.1 Hematopoietic Stem Cells

Self-renewal is defined as the process by which one stem cell divides to produce one or more stem cells, thereby maintaining the undifferentiated state[4]. Stem cells can undergo self-renewal via symmetric or asymmetric division[5]. Symmetric self-renewal occurs when one stem cell divides to produce two stem cells, resulting in the expansion of the stem cell pool. Asymmetric self-renewal occurs when one stem cell divides to produce one stem cell and a more restricted progenitor, providing maintenance of the stem cell pool. Thus, when HSCs undergo asymmetric self-renewal, it maintains a constant number of HSCs while also giving rise to the mature and functional hematopoietic cells[6]. Many of the mature blood cells at the bottom of the hierarchy have limited lifespans, thus the balance of self-renewal and differentiation is critical in maintaining equilibrium of the hematopoietic system[7]. The disruption of this balance (i.e. abnormal self-renewal of progenitor cells or failure of progenitor cells to differentiate) can lead to severe disease. The advantage of the increasing proliferative capacity down the hierarchy is that it allows HSCs to be maintained in a relatively lower proliferative state. This decreased proliferative pressure minimizes the risk of mutations associated with DNA replication and cell division, thereby prolonging the life of these cells[3]. Furthermore, HSCs are mostly in the G_0 phase of the cell cycle—a relatively decreased metabolic phase—which protects them from mutagenic byproducts.

1.2 Acute Myeloid Leukemia

Acute myeloid leukemia (AML) is a deadly cancer of the blood and bone marrow (BM) defined by the accumulation of immature and non-functional myeloid progenitor cells (myeloblasts)[8,9]. Furthermore, there is a parallel decrease in the production of mature blood cells required for life-sustaining functions, ultimately affecting survival. The median age of diagnosis is 68 years, with

an overall annual incidence of 4.3 per 100,000 which increases with age. AML is the most common type of acute leukemia in adults and has the shortest 5-year survival rate of 24% with a median survival of 8.5 months[10]. Current cure rates of AML patients are 35-40% in adults aged 60 or younger, and 5-15% in patients older than 60 years of age[11]. AML is generally caused by numerous mutations that occur at the top of the hematopoietic hierarchy typically in HSCs or progenitors[12]. Previous research suggests a common pattern of mutations in genes involved in epigenetic regulation which are present in HSCs and occur early in AML disease progression[11]. When HSCs or myeloid progenitors acquire these mutations and rapidly proliferate, the remainder of the hierarchy is disrupted, leading to AML disease.

1.2.1 AML Prognosis

There are many prognostic factors in AML that can aid in predicting disease progression, survival, and deciding appropriate treatment options. These factors can be split categorized as patient-related and disease-related[11]. Patient-related prognostic factors include increasing age, pre-existing and co-existing conditions, history of myelodysplastic syndrome (MDS), and performance status. Disease-related factors include white blood cell count, prior cytotoxic therapy, and genetic changes in leukemic cells. Based on these prognostic factors, patients can be segregated into three groups: favourable, standard, and unfavourable[9]. Knowing the patients' classification of prognostic subgroup improves the efficacy of AML treatment. More recent research on AML prognostic criteria focuses on genetics where three markers are currently used in the clinic: *NPM1*, *CEBPA*, and *FLT3* mutations[11].

1.2.2. Heterogeneity in AML

Due to the variety of prognostic factors described, there is heterogeneity among AML patients (interpatient heterogeneity), making treatment very difficult. In addition to the various prognostic factors discussed, there is also heterogeneity at the cellular level (intra-patient heterogeneity)[9]. This encompasses the variability of cell commitment and differentiation in the hematopoietic hierarchy. This type of heterogeneity resulted in the development of different subgroups where each subgroup describes a morphologically distinct blast cell[9]. Another form of heterogeneity occurs at the molecular level, with one of the major factors being epigenetics. Epigenetic modifications are changes in gene expression that are not genetic[13]. Some epigenetic processes include DNA methylation, histone modification, and chromatin remodeling. These processes are vital in HSCs as they play a role in maintaining the undifferentiated state. When there are aberrations of epigenetic modifications in HSCs, it can disrupt normal hematopoiesis and contribute to the progression of AML[13,14]. Furthermore, it is also suggested that high post-therapy relapse rates are attributed to disease heterogeneity[13].

1.2.3 AML Treatment

Over the years, many treatment options have become available in the treatment of AML patients with each treatment entailing different intensities, limitations, and remission rates. For the past 50 years, the standard induction chemotherapy treatment for AML has been the "7+3" regimen. This treatment is where patients are infused with cytarabine for seven continuous days, coupled with an anthracycline on the first three days[14]. This treatment has been successful in achieving complete remission (CR1) in 60-85% of adults aged 60 or younger and 40-60% of adults aged 60 or older[11]. Although there are lower CR rates in older patients, they should not be

dismissed of receiving the treatment unless they exhibit specific adverse prognostic factors where the treatment will be less likely to benefit[11]. There are also a variety of other treatment options that include microenvironment targeting, cell cycle checkpoint inhibitors, and epigenetic therapies[14]. While AML is associated with a high success of chemotherapy-induced remission, it is also accompanied by rapid and high relapse rates with poor response to subsequent therapies. As a result, relapsed AML patients only have a 10% probability of long-term survival[15]. After patients achieve initial CR, it is important to follow up induction therapy with a form of postinduction therapy to prevent relapse[9]. One of the most effective postinduction therapies is hematopoietic stem cell transplantation (HSCT).

1.3 Hematopoietic Stem Cell Transplantation

There are two methods of HSCT: autologous and allogenic HSCT. Autologous transplantation involves using the patient's own stem cells to treat the disease. Advantages of this method of transplantation is that they are readily available and can be administered without the need of identifying an HLA-matched donor[16]. However, in AML, since the cancer originates in the blood and bone marrow, cells taken from the blood may contain tumour cells that could cause relapse, making the procedure ineffective. Research is being conducted to find ways to kill the contaminated cells of the donated sample ex vivo[9]. Therefore, currently, allogeneic HSCT is preferred in the treatment of AML where the source of the stem cells come from an HLA-matched donor[11]. The advantage of allogeneic transplantation is that the graft will not contain any residual tumour cells as it comes from a healthy donor. There is also a lower risk of relapse compared to autologous transplantation[16]. Another benefit of allogeneic HSCT is the graft-versus leukemia (GVL) effect, where the donor immune cells aid in eliminating any residual leukemic cells.

Therefore, GVL contributes to reducing the risk of relapse in AML patients[17]. However, the major limitation of allogenic HSCT is graft-versus-host disease (GVHD) where the donated stem cells view the recipient body as foreign and attacks the healthy host cells[18].

Prior to HSCT, some form of conditioning is required. Chemoradiotherapy is normally selected due to its antileukemia potency and immunosuppression to allow for adequate engraftment[11]. Another conditioning option is total-body irradiation (TBI). The purpose of both conditioning methods are to rid the body of proliferating leukemic cells via chemotherapy and inactivate any surviving leukemic cells via TBI[19]. After this is accomplished, the patient is transfused with the donor's stem cells which will eventually home to the bone marrow and repopulate the hematopoietic system with healthy blood cells. Tolerance to chemotherapy used in the conditioning regimens is patient dependent and therefore regimens are tailored to disease risk and patient fitness where lower risk fit patients receive nonmyeloablative regimens. Although older patients must receive reduced-intensity conditioning, it cannot be too reduced as it can leave behind too many residual leukemic cells, thereby increasing chances of relapse[11].

There are many limitations that are associated with HSCT that are important to consider before administering the treatment to patients. Allogeneic sources of stem cells may not be readily available for patients who do not have HLA-matched siblings. Although this affects all patients, it particularly affects minority groups where there is not an abundant source of HLA-matched donor samples[11]. Thus, these patients would have to settle for partially matched family or unrelated donors, which could decrease the efficacy of the transplant. Furthermore, there can be long-term complications that negatively impact quality of life. For example, GVHD can be a severe and debilitating clinical burden itself, diminishing the benefits of the transplant. All of these

complications are also accompanied by a cost of approximately $87,000-$300,000 per transplant[20], and the utilization of scarce healthcare resources.

1.4 Stemness in AML

a) Leukemic Stem Cells

Poor survival of adult AML patients has been recently attributed to stemness properties measured by the presence of leukemic stem cells (LSCs). It is suggested that when a HSC or progenitor cell acquires a mutation, it results in a new entity defined as the leukemic stem cell[21]. LSCs are biologically distinct from other cell types and possess the same characteristics that define regular stem cells as they can self-renew and are highly proliferative[21]. These mutated cells disrupt the hematopoietic hierarchy, leading to the accumulation of myeloblasts. In addition to self-renewal, LSCs have also shown drug resistance characteristics, suggesting that they contribute to AML relapse[22]. There are numerous ways to identify LSCs, one of which include cell surface markers. A common biomarker of LSCs include high CD34 expression and lack of CD38 expression (CD34+/CD38-)[22]. Leukemic cells possessing a CD34+/CD38- immunophenotype are generally capable of engrafting and producing AML disease in immunodeficient mice. This immunophenotype also has strong prognostic value where patients with higher expression of CD34+/CD38- expression had a median relapse-free survival of 5.6 months, compared to 16 months in patients with lower CD34+/CD38- expression[22]. However, there is heterogeneity among LSCs as previous research has demonstrated that even cells lacking the CD34+/CD38- phenotype were able to engraft in immunodeficient mice, suggesting that LSCs may not be restricted to a specific phenotype[23]. To address this inconsistency, a transcriptional "LSC17" signature for AML-LSCs was defined that has prognostic potential[24]. The LSC17 signature was developed by analyzing differentially expressed genes between LSC+ and LSC- cell fractions[25]. In addition to

being correlated with LSC activity, the LSC17 signature has prognostic potential as it has been demonstrated that higher LSC17 scores are correlated with poor clinical outcomes including shorter survival rates[24,25].

Although there are various ways to study LSC activity, to date, the xenotransplantation LSC functional assay is considered the gold standard[26]. However, there are limitations to this assay which include the labour, expense, and expertise required for *in vivo* transplantation assessment. Furthermore, large numbers of viable cells from patients are required for successful xenotransplantation. Another limitation is that LSCs can only be measured in the minority of AML patients (<40%)[27] and thus fail to capture the majority of patients that succumb to relapsed disease. This ultimately limits the direct utility of xenograft assays to aid in routine clinical decision making.

b) Leukemic Progenitor Cells

Leukemic progenitor cells (LPCs) are positioned lower in the hematopoietic hierarchy with decreased self-renewal compared to stem cells. As there is a buildup of immature progenitor cells in AML, these cells are relatively more abundant than LSCs, with a frequency of 1/100 mononuclear cells for progenitors compared to $1/10^6$ for LSCs[28]. This makes leukemic progenitors easier to target and quantitate. Although leukemic progenitors have a decreased self-renewal capacity compared to LSCs, similar to their healthy counterparts, they are highly proliferative. Thus, these cells can be grown *in vitro* and quantitated via the colony forming unit (CFU) assay. BM or peripheral blood (PB) samples from patients are grown *in vitro* in a semi-solid methylcellulose-based media supplemented with growth factors that support the survival and proliferation of myeloid progenitors. Each colony that forms originates from a single cell with

myeloid/erythroid developmental potential and signifies a progenitor cell[29]. Furthermore, in contrast to the LSC assay, the *in vitro* detection of LPCs allows for the faster quantitation of leukemic progenitors, and with less expense and complexity, thereby making it more adaptable for use in a clinical setting.

1.5 DNA Methylation

As discussed before, aberrations in epigenetic processes can contribute to AML progression as well as disease heterogeneity. An important epigenetic modification in AML is DNA methylation as many AML-inducing mutations occur in genes responsible for this process[13]. DNA methylation is a repressive marker where gene expression in inhibited by preventing transcription factors from binding to the DNA[30]. The process involves the transfer of a methyl group onto the C5 position of cytosine, catalyzed by enzymes called DNA methyltransferases (DNMTs)[30,31]. Not only is there great heterogeneity in DNA methylation within and across patients, but very little is known specifically in the context of AML, making it difficult to use as a prognostic measure[32]. It can be valuable if a DNA methylation signature is developed such that AML patients can be stratified into high and low risk groups, allowing for more personalized treatment. To develop such a signature requires high-throughput experimentation, which can be achieved by Illumina's Infinium Methylation BeadChips[33]. This technology combines the Infinium I and Infinium II methylation assays, where probes are designed to target specific CpG sites in a particular DNA sample. The Infinium I assay uses two probes for each CpG locus, where each probe is designed to either complementary pair with cytosine (methylated) or thymine (unmethylated). The Infinium II assay uses one probe and a single base extension will determine whether the site is methylated or not

depending on the base added: if an adenine base is added, the site is unmethylated (complementary to thymine) and if a guanine base is added, the site is methylated (complementary to cytosine)[33].

1.6 Study Rationale

Our group has evaluated over 100 AML patients and completed multivariate survival analysis that uniquely incorporated *in vivo* and *in vitro* biological assays of patient cells obtained at diagnosis. We demonstrate that *in vitro* leukemic progenitor content is the strongest prognostic factor for overall and disease-free survival in comparison to other measures of stemness. We also observed that AML progenitor content correlates to curative outcomes of HSCT. This bifurcation of AML patients suggests two distinct classes of AML disease, where HSCT distinctly benefits patients with abundant leukemic progenitors. Furthermore, despite the refinement of genetic risk classification over the years, progenitor assays provide prognostic value that is not captured by current genetic-based risk assessments. To date, no robust methodology is available to prospectively identify and distinguish AML patients that are more likely to benefit from HSCT, thereby avoiding unnecessary HSCT associated morbidity and mortality.

1.7 Objective and Hypothesis

The objective of this thesis is to develop and validate a simple and robust method that could be used in future prospective clinical studies as a prognostic indicator (biomarker) of AML patients that would most benefit from HSCT leading to long term survival and disease cure. *We hypothesize that leukemic progenitor content correlates to post-HSCT survival.* In order to test this hypothesis, this study was broken down into three specific aims:

1. Quantitate leukemic progenitor content to expand the preliminary observations into larger patient groups to determine prognostic value

2. Determine whether LPC content is associated with post-transplantation survival rates
3. Identify DNA methylation loci and/or gene expression signatures that distinguish LPC+ vs LPC- patients

2.0 Materials and Methods

2.1 Primary human AML patient samples

All primary AML patient samples were provided by the Juravinski Hospital and Cancer Centre in Hamilton, Ontario, and the London Health Sciences Centre (LHSC) at the University of Western Ontario in London, Ontario. All patient samples were collected upon the informed consent of each patient in compliance with the protocols established by the Research Ethics Board at McMaster University and the University of Western Ontario. The primary patient samples came from either PB apheresis or BM aspirates. For all patients, PB samples were prioritized and when PB samples were not available, BM samples were used. All samples that were used in this study were taken at diagnosis to ensure consistency across samples. Upon processing the donated patient samples, the cells were cryopreserved in liquid nitrogen tanks in vials consisting of 90% FBS base with 10% DMSO. The clinical characteristics of all AML patients used in this study are summarized in **Table 1**.

Table 1. Clinical characteristics of 121 patients tested for functional LSC and leukemic progenitor content

	All patients	LPC+	LPC-	LSC+	LSC-
n	121	68	48	52	57
Age					
years, median (range)	63 (21-94)	63 (24-94)	63 (21-89)	64 (24-89)	59 (21-94)
Sex					
% Female	43	40	46	44	42
Blasts					
%, median (range)	86.3 (29.8-97.7)	85.7 (29.8-97.7)	87.0 (55-97.2)	89.1 (34-97.7)	82.3 (29.8-97.6)
ELN*					
Favourable, *n*	17 (22%)	9 (19%)	8 (30%)	7 (23%)	9 (23%)
Intermediate, *n*	32 (42%)	17 (35%)	13 (48%)	11 (37%)	16 (41%)
Adverse, *n*	28 (36%)	22 (46%)	6 (22%)	12 (40%)	14 (36%)

*Information unavailable for one or more patients

2.2 Leukemic Progenitor Cell Assay

2.2.1 Overview of the LPC Assay

The LPC assay employs a similar protocol to the CFU assay that was previously discussed[29]. The purpose of the LPC assay is to quantify the frequency of LPCs in human AML patient samples. After thawing the cryopreserved AML patient samples using a 37ºC water bath, the cells were seeded in semisolid methylcellulose media (Methocult GF #H4434; Stemcell Technologies) into a 12-well non-tissue culture plate. When seeding the cells, an increasing gradient of cell seeding densities was used (ranging from 1000-50000 cells/well) and each sample was seeded as triplicates, which was in accordance with previous protocols established by our group[34]. By using triplicates for each seeding density, the technical variance becomes suppressed, thereby yielding more accurate measurements for colony frequencies. The cells were then incubated for 14 days, periodically checking for any contamination to ensure quality results.

Following the 14-day incubation period, colonies were either manually counted and/or imaged for automated colony counting using the Operetta High

Content Screening platform (Perkin Elmer). Subsequent counting of individual colonies was done using the custom image analysis scripts on the Acapella software. Leukemic colonies were counted using a criteria of minimum 15 cells per colony, which is similar to thresholds used by other groups[35]. The threshold of 15 cells per colony was established due to the nature of the detecting algorithms on the Acapella software. When using minimum values of less than 15, the software was unable to accurately identify colonies, altering the colony counts for various samples. Therefore, by using a consistent minimum threshold of 15 cells per colony, the methodology was normalized across both automatic and manual counting techniques. For each patient sample, LPC frequencies were calculated as a measure of LPCs/10,000 cells. Patients were then stratified into different categories based on LPC content using a threshold of 3 LPCs/10,000 cells. Patients with an LPC frequency of 3 or less LPCs/10,000 cells were classified as LPC-low (LPC-). Patients with an LPC frequency or greater than 3 LPCs/10,000 cells were classified as LPC-high (LPC+). An overview of the LPC assay is outlined in **Supplementary Figure 1**.

2.2.2 Reliability and Validity of the LPC Assay

Furthermore, our group has demonstrated that 90% or more of colonies arise from leukemic cells[36]. This ensures that the colonies observed from AML patient samples are leukemic in origin, and not signifying a healthy progenitor. The reliability of the customized image analysis scripting was confirmed by directly comparing manual counts versus automated counts of progenitor colonies. Both methods of detecting leukemic colonies *in vitro* were shown to match (**Supplementary Figure 2**). Not only does the automated method allow for the ability to study a

larger amount of AML patient samples in a given timeframe, but it allows it in a rapid and non-biased manner[36].

2.3 DNA Extraction

Following the thawing of cryopreserved cells of the primary AML patient samples, 1-2 million cells were isolated for DNA extraction. DNA extraction was completed using the protocol provided by Qiagen's DNeasy Blood & Tissue Kit. Bulk DNA from primary AML patient samples were extracted and stored in -80ºC in aliquots of 100µL.

2.4 DNA Methylation Analysis

2.4.1 Illumina Methylation BeadChips Assay

Extracted DNA from 67 primary AML patient samples (**Table 2**) were shipped to Dr. Bekim Sadikovic's lab at the LHSC. Dr. Sadikovic's group applied the Infinium® HumanMethylation450 BeadChip assay (described in **section 1.5**) on all 67 patient samples to obtain raw DNA methylation data. The raw data were then sent to our group for further analysis.

Table 2. Clinical characteristics of 67 patients used for DNA methylation analyses

Sample	Age	Sex	LPC+/-	LSC+/-	HSCT
A295	52	m	LPC-	LSC+	No
13814	74	m	LPC-	LSC-	No
14939	42	f	LPC+	LSC+	No
16158	69	f	LPC+	LSC+	No
16308	N/A	m	LPC-	LSC+	Yes
16406	52	m	LPC-	LSC+	No
18182	69	f	LPC-	LSC-	No
18359	69	f	LPC-	LSC-	No
27264	89	f	LPC-	LSC+	No

A151	67	m	LPC+	LSC+	No
A233	77	f	LPC-	LSC-	No
A422	65	f	LPC+	LSC+	No
A254	65	f	LPC+	LSC+	No
A374	78	f	LPC+	LSC+	No
A421	58	m	LPC-	LSC-	No
A375	79	m	LPC+	LSC+	No
A241	79	m	LPC+	LSC-	No
A449	55	f	LPC+	LSC-	No
A472	73	m	N/A	N/A	No
A494	71	f	LPC+	LSC-	Yes
A559_PB	64	m	N/A	LSC+	Yes
A559_BM	64	m	N/A	N/A	Yes
A521	59	m	N/A	N/A	No
16626	47	m	LPC+	LSC+	No
15355	50	f	LPC+	LSC-	No
22472	64	m	LPC+	LSC-	No
19481	74	m	LPC+	LSC-	No
A345	55	m	LPC-	LSC-	No
A413	60	m	LPC-	LSC+	Yes
13685A	N/A	m	LPC-	LSC-	N/A
A407	67	m	LPC+	LSC-	No
A242	84	m	LPC-	LSC-	No
22174	68	f	LPC+	LSC+	No
12489	59	m	LPC-	LSC-	No
A069	64	m	LPC+	LSC+	No
14170	61	m	LPC-	LSC-	No
10169A	N/A	N/A	N/A	LSC-	Yes
25897	66	m	LPC-	LSC-	No
18675	55	f	LPC+	LSC+	No
10535A	N/A	m	LPC-	LSC-	No
14419A	N/A	N/A	LPC-	LSC-	N/A
15224	57	m	LPC+	LSC-	No
29560	39	f	LPC+	LSC+	Yes
17126A	N/A	N/A	N/A	LSC+	N/A
A408	79	m	LPC+	LSC-	No
13065	84	m	LPC+	LSC-	No
A136	43	m	LPC-	LSC-	No
20476	74	f	LPC-	LSC+	No
A382	66	f	LPC-	LSC+	No

A439	34	f	LPC-	LSC-	Yes
18758	54	m	LPC-	LSC-	No
12274A	N/A	m	LPC-	LSC-	N/A
A403	45	f	LPC+	LSC+	No
18947	61	f	LPC+	LSC-	No
A357	32	m	LPC-	LSC-	No
17694	63	m	LPC+	LSC-	No
26550	70	m	LPC+	LSC-	No
A038	54	f	LPC-	LSC-	No
12526	31	m	LPC-	LSC-	No
A477	62	m	LPC-	LSC+	Yes
15407	79	m	LPC-	LSC-	No
A379	76	m	LPC+	LSC-	No
A366	27	f	LPC-	LSC-	No
A360	56	f	LPC+	LSC+	No
28787	65	m	LPC+	LSC+	No
A499	69	f	LPC-	LSC-	N/A
A002	55	m	LPC-	LSC-	No

2.4.2 Processing and analysis of methylation raw data

The raw DNA methylation data were provided in the form of Intensity Data (IDAT) files by Dr. Sadikovic's group. The data were processed and analyzed using a workflow demonstrated by Maksimovic et al., using several Bioconductor packages in R studio[37]. The version of R studio used at the time of analysis was v2023.03.0. Prior to beginning the analysis, the raw IDAT files for all 67 AML patient samples were loaded into the environment. The first step in processing the data was to apply quality control methods. The detection p-values for all probes across all samples were calculated. Poor quality samples were removed from the analysis, such as those with a detection p-value of greater than 0.05. The next step was normalization, which minimizes the unwanted variation, such as differences in signaling, both within and between samples[37]. While there are many different normalization methods, the accepted method for analyzing global methylation differences in human cancers is functional normalization[38]. Therefore, the functional

normalization method was applied to this dataset. The last step in processing the data was filtering, where specific probes were removed including poor performing probes (detection p-value < 0.01), probes from sex chromosomes, and cross-reactive probes.

To analyze the data, a linear model was employed by creating a contrast matrix. The matrix included LPCs (LPC-high versus LPC-low) as the factor of interest, and sex to control for the individual variation. Upon fitting the contrast matrix into the linear model, the number of differentially methylated CpGs for each comparison (LPC+ versus LPC- and male versus female) were obtained. As there are hundreds of thousands of tests being performed, it is important to adjust the p-value for multiple testing[37]. Therefore, the Benjamini and Hochberg's (BH) adjustment method was used, which employs a 5% false discovery rate (FDR)[39,40].

2.5 Statistical Analyses

Prism 6 (v5.0a; GraphPad) software was used to plot survival curves and calculate median survival times. All univariate survival analyses were performed using the Kaplan-Meir method via Prism 6 software (v5.0a; GraphPad). The Mantel-Cox statistical test was used to calculate significant differences in survival between different patient groups. All multivariate survival analyses were performed using Cox regression. The DNA metylation analyses were performed using R studio (v2023.03.0) Specific R studio packages used include: limma (v3.54.2), minfi (v1.44.0), knitr (v1.42), IlluminaHumanMethylation450kanno.ilmn12.hg19 (v0.6.1), IlluminaHumanMethylation450kmanifest (v0.4.0), and missMethyl (v1.32.1).

3.0 Results

3.1 LPCs serves as a prognostic predictor of AML patient survival

3.1.1 Leukemic progenitor content emerges as a robust predictor of AML patient outcome

To date, there have not been studies that comprehensively measure leukemic stemness across the same set of patients. In order to determine which functional and molecular measures of leukemic stemness correlate with each other, we compared the clinical relevance of four different measures of stemness in a set of 21 patients. The four measures were: CD34 expression, LSC17 signature, functionally defined LSC presence, and leukemic progenitor cells (LPCs). The two molecular measures of stemness measured via flow cytometry and rapid laboratory tests (CD34 and LSC17, respectively) do not show as clear of a correlation with overall survival (OS) when compared to the two functional measures (LSC and LPC) (**Figure 3.1**). Of the four parameters, the LPC assay showed the strongest association with overall survival in our patient subset and therefore, we expanded the data to a larger set of 121 patients.

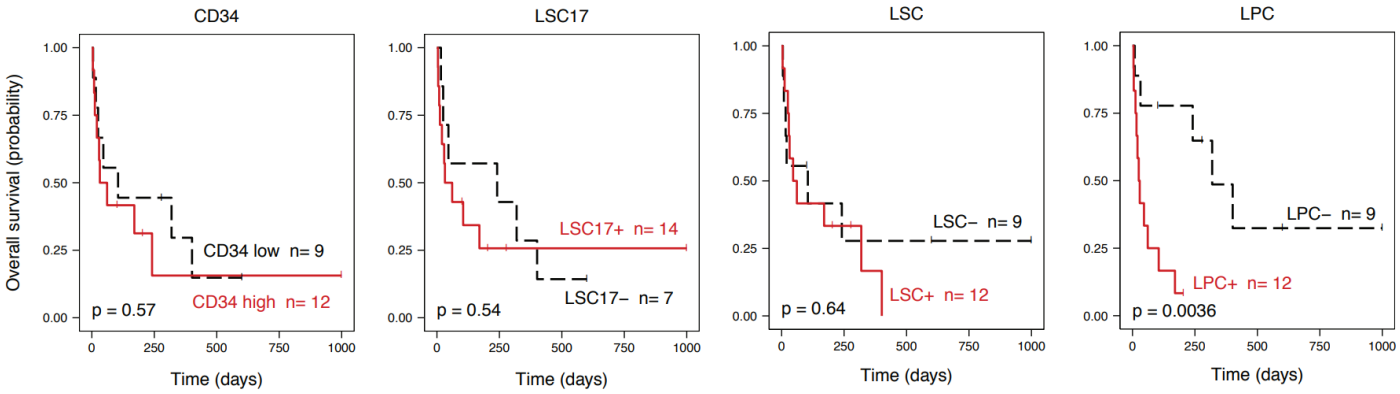


Figure 3.1: **Substantial variation exists between different functional and molecular measures of leukemic stemness.** All four panels display Kaplan-Meier estimates of Overall Survival based on CD34 expression (flow cytometry), LSC17 score (Nanostring assay), LSC content (xenograft assay), or LPC content (CFU assay).

From this cohort of 121 patients, a total of 111 were tested for LSCs in immune deficient xenograft models. The sample size used here is higher than any previous study for functional in

vivo LSC evaluation to date, which have only ranged from 25 to 44 patients[41–44]. With the increased the sample size, we were still able to produce results that were consistent with these previous studies, demonstrating that patients with detectable LSCs (LSC+) had shorter Overall Survival (OS) compared to those without LSCs (LSC-) (**Figure 3.2A**).

Our total patient cohort of 111 patients that were tested for LSCs included patients under palliative care, patients that were only treated with standard chemotherapy, and patients that were treated with both standard chemotherapy followed by hematopoietic stem cell transplantation (HSCT). HSCT is the most effective therapy for durable remission for AML patients, thus having a profound impact on survival outcomes. In order to truly demonstrate the negative prognostic impact of LSCs on AML patient survival, we repeated the analysis with a subgroup of patients that were more uniform with respect to therapeutic treatment and intent. This subgroup included patients that were only treated with high dose chemotherapy, thus removing palliatively-treated patients and patients who received HSCT. Transplanted patients have already been shown to increase overall survival; therefore, to control for the profound effect of HSCT on survival, we removed these patients in order to truly determine the prognostic value of the variable under study. When examining this new subgroup, the presence of LSCs no longer had a significant negative impact on OS (**Figure 3.2A**).

From the same pool of 121 patients, we next used the LPC assay for the efficient detection and quantification of leukemic progenitors *in vitro* in a total of 116 patients. We stratified the pool of patients into the same two subgroups used when evaluating for LSCs: 1) All patients and 2) patients that were only treated with high dose chemotherapy, while excluding palliatively-treated

patients and patients who received HSCT. Similar to *in vivo* LSC detection, patients with large amounts of leukemic progenitors (LPC+) displayed shorter OS when compared to patients with lower amounts of leukemic progenitors (LPC-) (**Figure 3.2B**). However, in contrast to LSCs, OS was shorter for patients with higher amounts of leukemic progenitors and this effect was more significant in patients treated with standard induction chemotherapy alone.

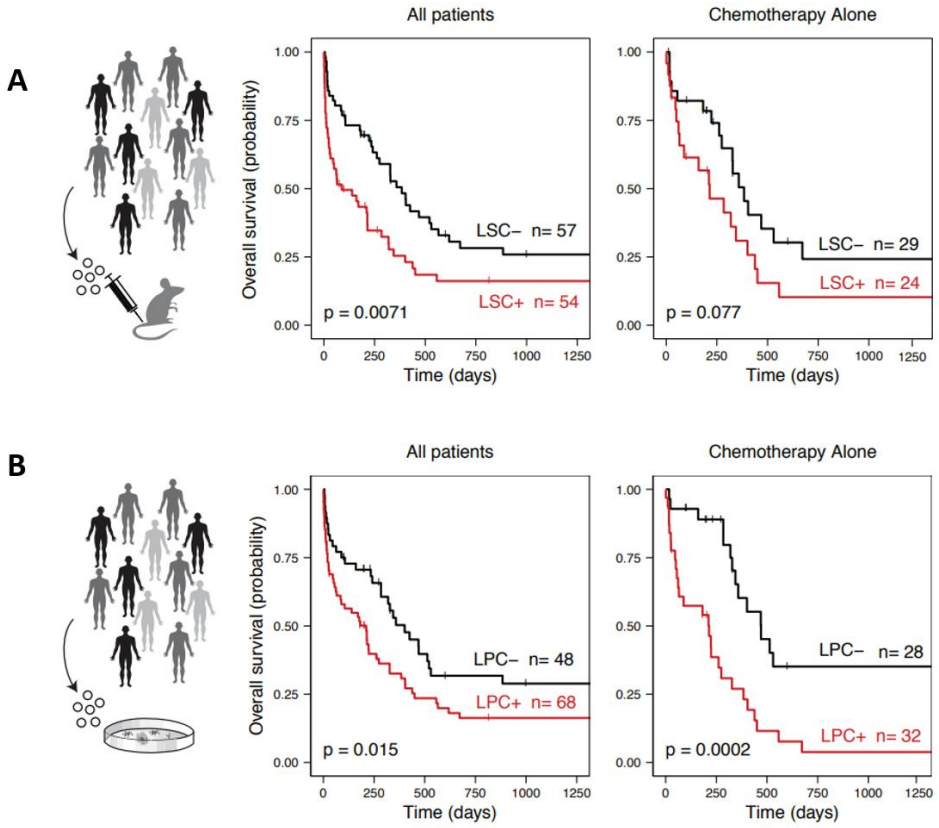


Figure 3.2: AML patient survival based on functional LSC and LPC content. A) Kaplan-Meier estimates of Overall Survival in LSC- and LSC+ patient subsets. Plots display all 111 patients tested in xenograft assays (left), or a subset of these patients treated with high intensity chemotherapy alone (palliative and HSCT-treated patients excluded; right). **B)** Kaplan-Meier estimates of Overall Survival in LPC- and LPC+ patient subsets. Plots display all 116 patients tested in colony forming LPC assays (left), or a subset of these patients treated with high intensity chemotherapy alone (palliative and HSCT-treated patients excluded; right).

3.1.2 Leukemic progenitor content represented an independent prognostic factor when compared to clinical guidelines of AML patient risk stratification

From the analysis discussed earlier, progenitor detection was achieved robustly in 58% of total AML patients tested, whereas LSC detection was only achieved in 48% of total patients. Therefore, to further investigate the performance of progenitor detection, we tested it against additional prognostic indicators currently used in the clinic for AML patients.

The most commonly used prognostic evaluation for AML patients is the European Leukemia Net (ELN) 2017 guidelines which stratifies patients into risk categories based on cytogenetics and mutational status[45]. Thus, we created a Multivariate Cox Proportional Hazards model of OS including ELN risk stratification, LSC engraftment *in vivo*, and LPC detection *in vitro* (**Figure 3.3A**). Leukemic progenitors showed to be a strong independent prognostic factor in predicting OS when included as a covariate. However, in contrast, LSCs were unable to retain prognostic power as they had no significant impact on OS when coupled with covariates of LPC content and ELN stratification.

Next, it was important to evaluate the impact of LPCs and LSCs on event-free survival (EFS) because while OS includes all mortality events, regardless of whether they are attributable to disease or not, EFS is specific to disease progression/persistence. To do this, we established a patient subgroup where all patients were treated with induction chemotherapy, and did not receive HSCT. This subgroup was selected due to the profound effects of chemotherapy and HSCT on disease status and outcomes. Both the presence of LSCs and LPCs had a significant negative impact on EFS, although the LPC-associated effects on EFS were stronger (**Figure 3.3B-C**).

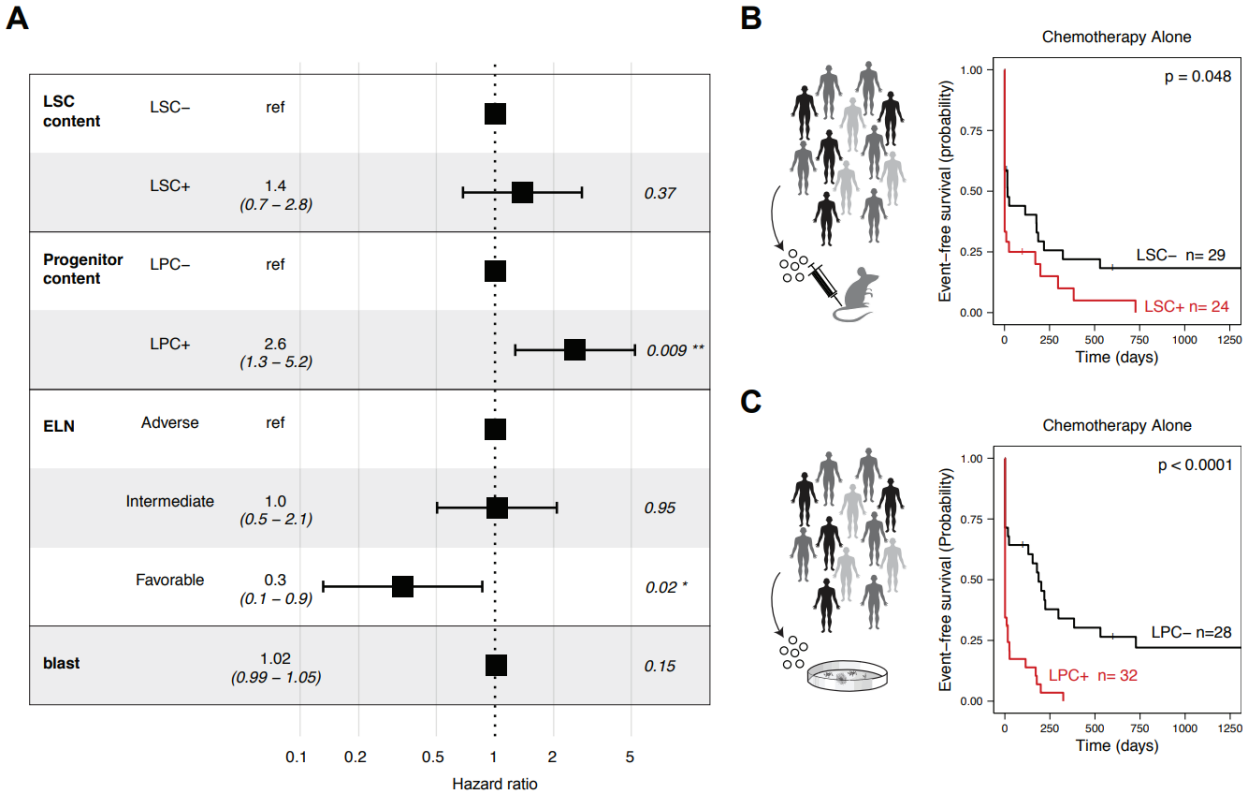


Figure 3.3: Functional LPC frequencies represent an independent prognostic factor in human AML. A) Forest plot showing multivariate analysis of Overall Survival in AML patients treated with high intensity chemotherapy only. Unadjusted HRs (squares) and 95% confidence intervals (horizontal lines) are shown. **B)** Kaplan-Meier estimates of Event-free Survival in LSC- and LSC+ patient subsets. **C)** Kaplan-Meier estimates of Event-free Survival in LPC- and LPC+ patient subsets.

3.2 Leukemic progenitor content correlates with post-HSCT survival

3.2.1 LPC+ vs LPC– Survival Outcomes Post-HSCT

From the total patient cohort of 121 patients, we were able to receive clinical annotation regarding presence and/or absence of HSCT for 103 patients. HSCT significantly improved OS in both LPC+ and LPC- patients, but LPC+ patients appear to benefit more strongly than LPC- patients (**Figure 3.4**). The benefit of HSCT for LPC+ patients can be more clearly understood

when visualizing the median survival times for each patient category. The median survival for LPC+ patients that were transplanted and not transplanted was 1231 days and 182 days, respectively. Additionally, the median survival for LPC– patients that were transplanted and not transplanted was 1071 and 344 days, respectively. Here, we see an approximate 7-fold increase in survival time when LPC+ patients are transplanted, whereas only an approximate 3-fold increase for LPC– patients, demonstrating the increased benefit of HSCT for LPC+ patients.

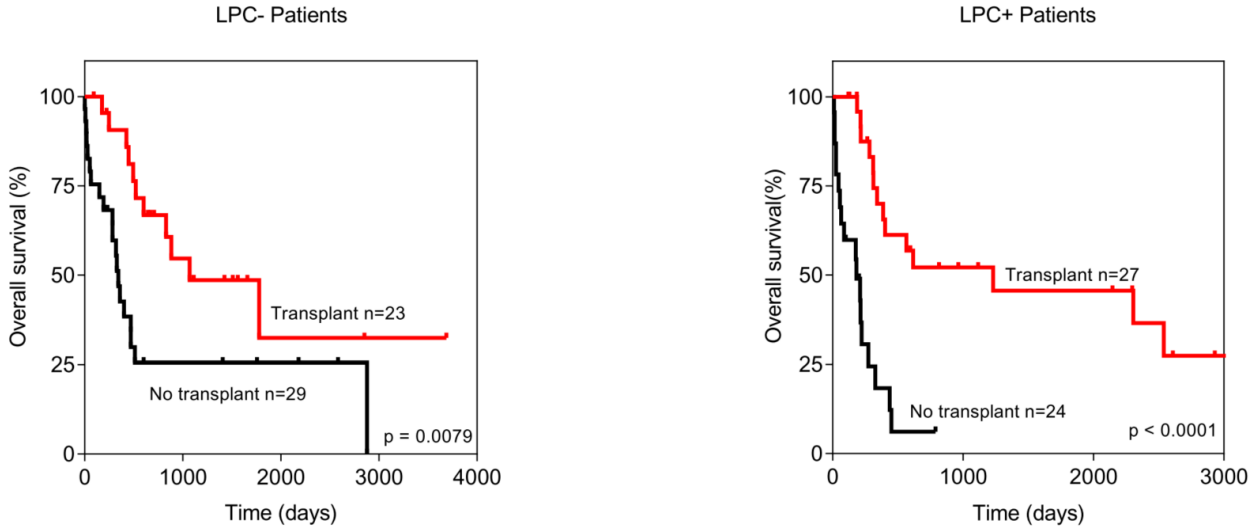


Figure 3.4: LPC content measured by LPC assay predicts transplantation benefit in AML patients. Kaplan-Meier estimates of Overall Survival in transplanted vs. non-transplanted patients, stratified by LPC content.

3.3 No identifiable DNA methylation loci distinguishing LPC+ and LPC- patients

Using the methods discussed in *Section 2.4.2*, there were 0 significantly differentially methylated CpGs when comparing LPC+ versus LPC- patients, based on a 5% FDR. The total number of CpG sites tested were 797815. To confirm the quality of the raw data obtained, we repeated the same analysis for sex, which resulted in a total of 3742 significantly differentially methylated CpG sites.

4.0 Discussion and Conclusion

Our results show that when comprehensively comparing different measures of cancer stemness, functional measures (i.e. LSC and LPC content) were superior in predicting OS compared to the phenotypic measures (i.e. LSC17 score and CD34 expression) (**Figure 3.1**). This suggests that primitive cell phenotypes often imply the presence of self-renewing leukemic cells even when they may not be functionally detectable in the sample, therefore emphasizing the importance of functional assays. Furthermore, these results suggest that a single surrogate phenotype or signature may not be sufficient enough to serve as a universal biomarker of leukemic stemness. This is consistent with findings from other groups where the transcriptional profiles of functionally defined LSCs actually differ based on CD34 expression[46]. While LSCs are typically defined by CD34+ expression, AML patients that express the NPM1 mutation completely lack CD34 expression[35], yet they still possess LSCs which appear to be biologically distinct from the rest of the leukemic cells. Therefore, it is important that functional assays are prioritized in order to accurately measure cancer stemness in patients.

Our group has demonstrated through our analysis of LSC content in relation to OS that functionally defined LSCs are a negative prognostic factor in adult AML. We accomplished this with the largest patient cohort to date, while continuing to reproduce results consistent with previous studies[41–44]. Although LSCs no longer significantly affected OS when controlling for therapeutic treatment within the patient cohort (**Figure 3.2A**), we have shown that high LSC content significantly decreases EFS in patients that were only treated with standard chemotherapy (**Figure 3.3B-C**). While this supports the clinical value of xenograft-defined functional LSCs, the variability in the results for OS when tested against the clinically utilized ELN guidelines suggests

that a more consistent measure of cancer stemness is required when predicting AML patient survival.

In contrast to LSCs, leukemic progenitor content was a strong predictor of OS under all conditions, including the therapeutically-controlled subgroup of chemotherapy-treated patients only (**Figure 3.2B**), and when tested against ELN risk stratifications in the multivariate analysis (**Figure 3.3A**). While previous studies have shown the correlation between LPC content and survival[47,48], none have tested against genetic risk measures. Therefore, our study has uniquely demonstrated that leukemic progenitor content has a stronger correlation with OS than that of the predominantly genetic-based risk assessment outlined by the ELN risk stratification. This highlights the comparatively more powerful prognostic utility of LPCs to predict patient survival outcomes. While we are aware that there is a 2022 version of the ELN guidelines, most of the patient samples from our cohort only had clinical annotations that were in accordance with the 2017 version. Therefore, for future studies, we will ensure to repeat the analyses with the updated ELN classifications. The clinical applicability of LPCs that we illustrated in our study provides a basis to extend our findings to other cancers in which the respective cancer stem cells (CSCs) have been detected *in vivo*. In addition to progenitors being more amendable to *in vitro* detection due to the relatively abundant frequency compared to stem cells[28], the considerably lower cell numbers needed for progenitor assays make it a more feasible diagnostic option in the clinic. Therefore, the LPC assay can be used for additional applications in AML research such as drug screening and in clinical trials where it is difficult to obtain large number of cells from patients. Furthermore, the use of progenitor content as an alternative to the CSC counterparts of different cancers represents a potential paradigm shift in the field of CSC research.

We demonstrated that patients with a higher amount of LPCs appear to better respond to HSCT as shown by the increase in OS compared to patients with lower amount of LPCs. (**Figure 3.4**). Therefore, this supports our hypothesis that leukemic progenitor content correlates to post-HSCT survival. Thus, this provides an initial basis to suggest that HSCT can offset the negative prognostic impact associated with high leukemic stem/progenitor cell content.

However, it is important to consider the confounding variables that can affect AML patient survival. These factors will aid in understanding the clinical biases and potentially result in further stratification based on certain categories. Therefore, with the collaboration of Dr. Xenocostas at the LHSC, we conducted a deep dive on a subset of the patient cohort from our HSCT survival studies. We focused on two specific clinical annotations: 1) The cause of death for all deceased patients and 2) The reason for non-transplanted patients not receiving a transplant. We received a summary of the clinical annotations across the four patient groups that were studied: LPC+ transplant, LPC+ no transplant, LPC– transplant, and LPC– no transplant (**Supplementary Table 1**).

Among transplanted patients (both LPC+ and LPC–), the most common reason for death was disease relapse (**Supplementary Table 1**). Although disease relapse is a prevalent clinical challenge to tackle, this shows the advances in allogeneic HSCT over the years. Previously, transplanted-related mortality was the primary cause of death and treatment failure among AML patients, which has now been replaced by disease relapse[49]. Additional reasons for death among transplanted patients included GVHD, graft failure, and multiorgan failure. The clinical

annotations for the non-transplanted patients showed greater variability. Among the non-transplanted patients, the most common reasons for death were treatment-related complications and disease relapse. Furthermore, there were many reasons for these patients to not get transplanted, some including death during admission for induction chemotherapy, failure to achieve remission, and age. This shows that the majority of the non-transplanted patients in our cohort would have otherwise been deemed ineligible for HSCT. Therefore, future studies investigating the progenitor-associated outcomes on post-HSCT survival should ensure that all patients in the cohort have received standard chemotherapy, achieved CR1, and are eligible for HSCT. With this newly defined patient cohort, replicating our original findings that LPC+ patients survive longer than LPC- patients following HSCT would allow us to more confidently conclude that LPCs do in fact play a major role in the positive outcomes of HSCT. As a result, the LPC assay would emerge as an immensely valuable tool in the clinic when determining transplant eligibility among AML patients.

Given the new clinical annotations (**Supplementary Table 1**), this provides insight into the potential for further patient stratification. For example, patients can be stratified based on a common reason for death (i.e. disease relapse), followed by subsequent survival analysis among this specific subset of patients. However, in order to do this, a much larger patient cohort would be needed, along with a balanced sample number under each sub-category.

Although we were not able to discover DNA methylation patterns distinguishing LPC+ and LPC- patients, this does not necessarily mean that there is no biological signature unique to LPCs. Further studies need to be conducted, either with a larger patient cohort for methylation

studies, or transition into investigations at the transcript level. Furthermore, the current methods need to be repeated with a more liberal FDR of either 10 or 20%, which can potentially yield significantly methylated sites. If the new FDR results in significantly methylated regions, this would provide an initial starting point to suggest which genes should be further analyzed at the transcriptional level, leading to future studies involving next-generation sequencing methods. Due to the relatively low frequency of LPCs present in patient samples, coupled with the heterogeneity across AML patients, it may be difficult to capture global differences in DNA methylation for a specific and rare cell type.

The conclusions of this study give insight into a new approach that should be taken for future studies investigating the biology of LPCs. Despite LSCs being less abundant than LPCs, researchers were able to develop a unique signature (i.e. the LSC 17 score) that distinguishes AML patients based on LSC content. While we only tested differential methylation, we were unable to develop a signature separating patients according to LPC content. This raises the question on whether LPCs detected in the LPC assay are truly representative of their position in the hematopoietic hierarchy. Currently, there have been no studies that have transplanted progenitor cells into mice to determine whether they have the ability to recapitulate AML disease. Future directions for our work will include xenotransplantation experiments by transplanting progenitor cells obtained from the LPC assay into immunodeficient mice. Outcomes from these experiments will be pivotal in the understanding of the biology of leukemic progenitors in AML.

Bibliography

1. Rieger, M. A. & Schroeder, T. Hematopoiesis. *Cold Spring Harb. Perspect. Biol.* **4**, a008250 (2012).

2. Jackson, J. D. Chapter 13 - Hematopoietic Stem Cell Properties, Markers, and Therapeutics. in *Principles of Regenerative Medicine (Third Edition)* (eds. Atala, A., Lanza, R., Mikos, A. G. & Nerem, R.) 191–204 (Academic Press, 2019). doi:10.1016/B978-0-12-809880-6.00013-8.

3. Bryder, D., Rossi, D. J. & Weissman, I. L. Hematopoietic Stem Cells. *Am. J. Pathol.* **169**, 338–346 (2006).

4. He, S., Nakada, D. & Morrison, S. J. Mechanisms of stem cell self-renewal. *Annu. Rev. Cell Dev. Biol.* **25**, 377–406 (2009).

5. Molofsky, A. V., Pardal, R. & Morrison, S. J. Diverse mechanisms regulate stem cell self-renewal. *Curr. Opin. Cell Biol.* **16**, 700–707 (2004).

6. Bujko, K., Kucia, M., Ratajczak, J. & Ratajczak, M. Z. Hematopoietic Stem and Progenitor Cells (HSPCs). in *Stem Cells: Therapeutic Applications* (ed. Ratajczak, M. Z.) 49–77 (Springer International Publishing, 2019). doi:10.1007/978-3-030-31206-0_3.

7. Seita, J. & Weissman, I. L. Hematopoietic Stem Cell: Self-renewal versus Differentiation. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2**, 640–653 (2010).

8. Acute Myeloid Leukemia Treatment (PDQ®)–Patient Version - National Cancer Institute. https://www.cancer.gov/types/leukemia/patient/adult-aml-treatment-pdq (2022).

9. Lowenberg, B., Downing, J. R. & Burnett, A. Acute Myeloid Leukemia. *N. Engl. J. Med.* **341**, 1051–1062 (1999).

10. Shallis, R. M., Wang, R., Davidoff, A., Ma, X. & Zeidan, A. M. Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. *Blood Rev.* **36**, 70–87 (2019).

11. Döhner, H., Weisdorf, D. J. & Bloomfield, C. D. Acute Myeloid Leukemia. *N. Engl. J. Med.* **373**, 1136–1152 (2015).

12. Acute Myeloid Leukemia (AML) - Hematology and Oncology. *Merck Manuals Professional Edition* https://www.merckmanuals.com/en-ca/professional/hematology-and-oncology/leukemias/acute-myeloid-leukemia-aml.

13. Goldman, S. L. *et al.* Epigenetic Modifications in Acute Myeloid Leukemia: Prognosis, Treatment, and Heterogeneity. *Front. Genet.* **10**, (2019).

14. Winer, E. S. & Stone, R. M. Novel therapy in Acute myeloid leukemia (AML): moving toward targeted approaches: *Ther. Adv. Hematol.* (2019) doi:10.1177/2040620719860645.

15. Ganzel, C. *et al.* Very poor long-term survival in past and more recent studies for relapsed AML patients: The ECOG-ACRIN experience. *Am. J. Hematol.* (2018) doi:10.1002/ajh.25162.

16. Champlin, R. Selection of Autologous or Allogeneic Transplantation. *Holl.-Frei Cancer Med. 6th Ed.* (2003).

17. Imamura, M., Hashino, S. & Tanaka, J. Graft-versus-leukemia effect and its clinical implications. *Leuk. Lymphoma* **23**, 477–492 (1996).

18. Ferrara, J. L., Levine, J. E., Reddy, P. & Holler, E. Graft-versus-host disease. *The Lancet* **373**, 1550–1561 (2009).

19. Quast, U. Total body irradiation—review of treatment techniquesin Europe. *Radiother. Oncol.* **9**, 91–106 (1987).

20. Broder, M. S. *et al.* The Cost of Hematopoietic Stem-Cell Transplantation in the United States. *Am. Health Drug Benefits* **10**, 366–374 (2017).

21. Jordan, C. T. The leukemic stem cell. *Best Pract. Res. Clin. Haematol.* **20**, 13–18 (2007).

22. Hanekamp, D., Cloos, J. & Schuurhuis, G. J. Leukemic stem cells: identification and clinical application. *Int. J. Hematol.* **105**, 549–557 (2017).

23. Sarry, J.-E. *et al.* Human acute myelogenous leukemia stem cells are rare and heterogeneous when assayed in NOD/SCID/IL2Rγc-deficient mice. *J. Clin. Invest.* **121**, 384–395 (2011).

24. Richard-Carpentier, G. & Sauvageau, G. Bringing a Leukemic Stem Cell Gene Signature into Clinics: Are We There Yet? *Cell Stem Cell* **20**, 300–301 (2017).

25. Ng, S. W. K. *et al.* A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature* **540**, 433–437 (2016).

26. Griessinger, E. & Andreeff, M. NSG-S mice for acute myeloid leukemia, yes. For myelodysplastic syndrome, no. *Haematologica* **103**, 921–923 (2018).

27. Shlush, L. I. *et al.* Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* **506**, 328–333 (2014).

28. Warner, J. K., Wang, J. C. Y., Hope, K. J., Jin, L. & Dick, J. E. Concepts of human leukemic development. *Oncogene* **23**, 7164–7177 (2004).

29. Sarma, N. J., Takeda, A. & Yaseen, N. R. Colony Forming Cell (CFC) Assay for Human Hematopoietic Cells. *J. Vis. Exp. JoVE* 2195 (2010) doi:10.3791/2195.

30. Moore, L. D., Le, T. & Fan, G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38**, 23–38 (2013).

31. Jin, B., Li, Y. & Robertson, K. D. DNA Methylation. *Genes Cancer* **2**, 607–617 (2011).

32. Li, S., Mason, C. & Melnick, A. Genetic and Epigenetic Heterogeneity in Acute Myeloid Leukemia. *Curr. Opin. Genet. Dev.* **36**, 100–106 (2016).

33. Infinium HumanMethylation450 BeadChip.

34. Boyd, A. L. *et al.* Identification of Chemotherapy-Induced Leukemic-Regenerating Cells Reveals a Transient Vulnerability of Human AML Recurrence. *Cancer Cell* **34**, 483-498.e5 (2018).

35. Vergez, F. *et al.* Phenotypically-defined stages of leukemia arrest predict main driver mutations subgroups, and outcome in acute myeloid leukemia. *Blood Cancer J.* **12**, 1–11 (2022).

36. Boyd, A. L. *et al.* Leukemic progenitor compartment serves as a prognostic measure of cancer stemness in patients with acute myeloid leukemia. *Cell Rep. Med.* **4**, 101108 (2023).

37. Maksimovic, J., Phipson, B. & Oshlack, A. A cross-package Bioconductor workflow for analysing methylation array data. *F1000Research* **5**, 1281 (2016).

38. Fortin, J.-P. *et al.* Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol.* **15**, 503 (2014).

39. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

40. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).

41. Pearce, D. J. *et al.* AML engraftment in the NOD/SCID assay reflects the outcome of AML: implications for our understanding of the heterogeneity of AML. *Blood* **107**, 1166–1173 (2006).

42. Pabst, C. *et al.* GPR56 identifies primary human acute myeloid leukemia cells with high repopulating potential in vivo. *Blood* **127**, 2018–2027 (2016).

43. Vargaftig, J. *et al.* Frequency of leukemic initiating cells does not depend on the xenotransplantation model used. *Leukemia* **26**, 858–860 (2012).

44. Griessinger, E. *et al.* Acute myeloid leukemia xenograft success prediction: Saving time. *Exp. Hematol.* **59**, 66-71.e4 (2018).

45. Döhner, H. *et al.* Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* **129**, 424–447 (2017).

46. Quek, L. *et al.* Genetically distinct leukemic stem cells in human CD34- acute myeloid leukemia are arrested at a hemopoietic precursor-like stage. *J. Exp. Med.* **213**, 1513–1535 (2016).

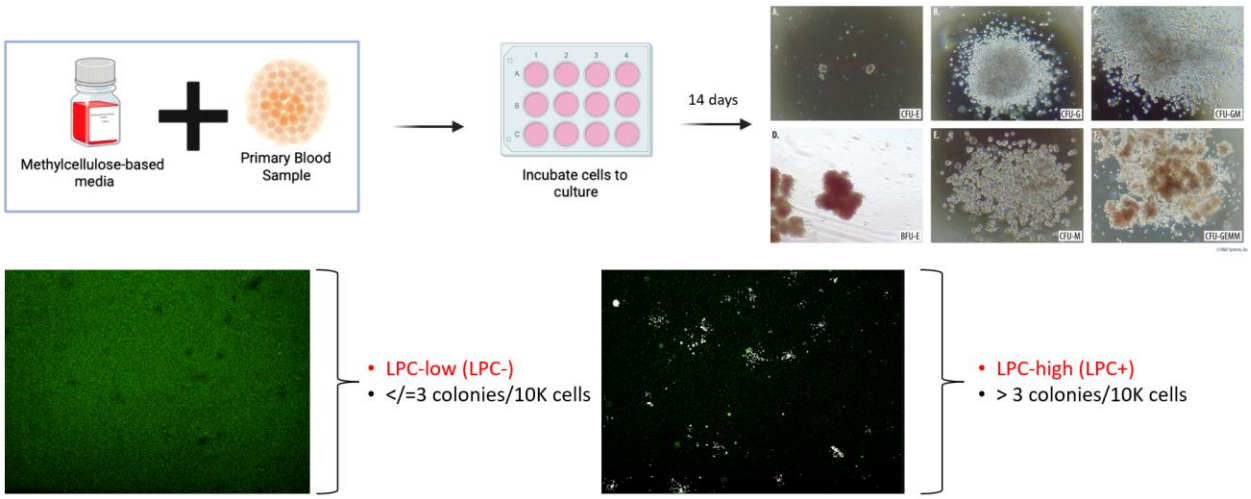
47. Delmer, A. *et al.* Multivariate analysis of prognostic factors in acute myeloid leukemia: value of clonogenic leukemic cell properties. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **7**, 738–746 (1989).

48. Sutherland, H., Blair, A., Vercauteren, S. & Zapf, R. Detection and clinical significance of human acute myeloid leukaemia progenitors capable of long-term proliferation in vitro. *Br. J. Haematol.* **114**, 296–306 (2001).

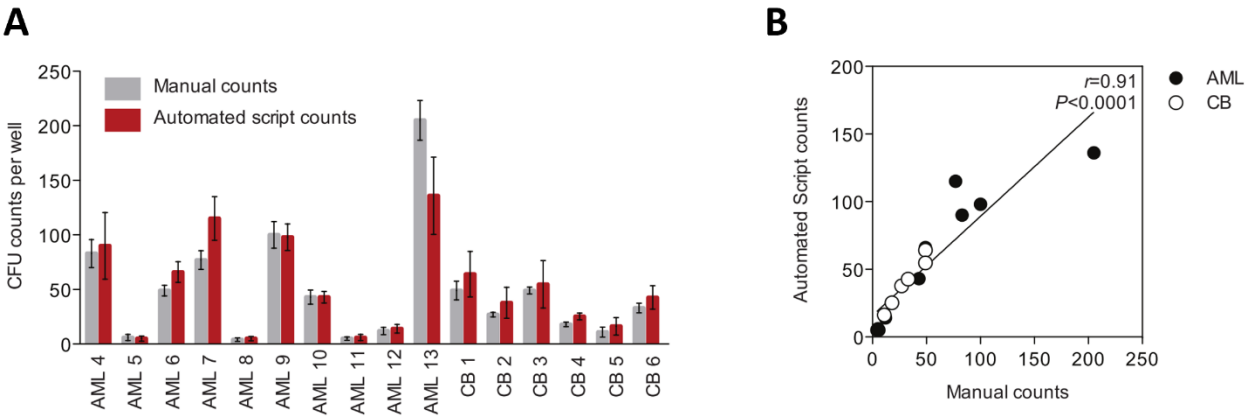
49. Webster, J. A., Luznik, L. & Gojo, I. Treatment of AML Relapse After Allo-HCT. *Front. Oncol.* **11**, (2021).

Supplementary Data

Supplementary Figures



Supplementary Figure 1: Overview of the LPC Assay



Supplementary Figure 2: Miniaturization and custom image analysis scripts allow high-throughput CFU testing. A) Automated script counts vs. manual counts of CFUs plated in 96-well platform. CB, healthy cord blood. **B)** Correlation of Automated script counts vs. Manual counts of CFUs plated in 96-well platform. CB, healthy cord blood.

Supplementary Tables

Supplementary Table 1. Clinical annotation of patients used for HSCT survival studies

Transplant Status	LPC+/ LPC-	Patient ID	Mortality Status	Reason for No Transplant	Reason for Death
Transplant	LPC+	A312	Alive	N/A	N/A
		A492	Alive	N/A	N/A
		A009	Deceased	N/A	No electronic record
		A372	Deceased	N/A	Disease relapse
		A457	Deceased	N/A	Disease relapse
		A494	Deceased	N/A	Disease relapse
		A409	Deceased	N/A	Disease relapse
		A430	Deceased	N/A	GVHD: GI tract & liver
		A577	Deceased	N/A	Disease relapse
	LPC-	A157	Alive	N/A	N/A
		A423	Alive	N/A	N/A
		A591	Alive	N/A	N/A
		A477	Deceased	N/A	Disease relapse
		A347	Deceased	N/A	Possible graft failure
		A537	Deceased	N/A	Multiorgan failure with likely underlying sepsis; unclear if patient relapsed.
		A485	Deceased	N/A	Progressive bowel ischemia following surgery for high grade small bowel obstruction
		A088	Deceased	N/A	Disease relapse
		A159	Deceased	N/A	Disease relapse
		A418	Deceased	N/A	Disease relapse
No Transplant	LPC+	A151	Deceased	Passed away during admission for induction chemotherapy	Early treatment related complications: Acute respiratory distress during induction chemotherapy
		A165	Deceased	Failed to achieve remission	Early treatment related complications: Multi-organ failure secondary to AML
		A254	Deceased	Passed away during admission for induction chemotherapy	Early treatment related complications: Fibroproliferative ARDS and multi-organ failure

		A449	Deceased	Passed away during admission for induction chemotherapy	Early treatment related complications: Intracranial hemorrhage
		A406	Deceased	Age? (Patient was 77, no discussion of transplant in clinic notes)	Disease relapse
		A407	Deceased	Passed away during admission for induction chemotherapy	Early treatment related complications: myocardial infarction, febrile neutropenia, and sepsis
	LPC-	A136	Alive	Good prognosis karyotype [t(8:21)]?	N/A
		A357	Deceased	Passed away during admission for induction chemotherapy	Early treatment related complications: ARDS and sepsis
		A382	Deceased	Failed to achieve remission	Disease progression
		A031	Deceased	No electronic record	No electronic record
		A295	Deceased	Lost to follow-up (Nov 2012)	No death summary - listed to follow-up (death confirmation based on obituary search)
		A302	Deceased	Passed away during admission for induction chemotherapy	Early treatment related complications: Neutropenic colitis & pneumonia
		A002	Deceased	No electronic record	No electronic record
		A121	Deceased	No indication in records for why patient wasn't transplanted	Disease relapse
		A421	Deceased	Patient declined out of country MUD transplant	Disease relapse
		A051	Unknown	No electronic record	No electronic record

		A038	Unknown	No electronic record	No electronic record
		A366	Unknown	Disease relapse before transplant	No record of death, but being treated palliatively