BLOOD AND MARROW
TRANSPLANT
CLINICAL TRIALS NETWORK

A Randomized, Phase II, Multicenter, Open Label Study Evaluating Sirolimus and Prednisone in Patients with Refined Minnesota Standard Risk, Ann Arbor I/2 Confirmed Acute Graft-Versus-Host Disease

BMT CTN PROTOCOL 1501
Version 2.0

Study Chairpersons
Joseph Pidala, MD, PhD[1] and Margaret MacMillan, MD[3]

Protocol Team

Amin Alousi, MD[2]
Claudio Anasetti, MD[1]
Joseph Antin, MD[4]
Peter Dawson, PhD[5]
Adam Mendizabal, PhD[3]
James Ferrara, MD[6]
Vincent Ho, MD[4]
Mary Horowitz, MD[7]
John Levine, MD[8]
Madan Jagasia, MD[10]
Lauren Kunz, PhD[9]
Eric Leifer, PhD[9]
Marco Mielcarek, MD[11]
Brent Logan, PhD[7]
Shernan Holtan, MD[3]
Javier Bolaños-Meade, MD[12]
Meludi Hamadani, MD[7]
Alan Howard, PhD[13]
Theresa Pritchard, MS[5]

Sponsored by the National Institutes of Health
National Heart, Lung, and Blood Institute
National Cancer Institute

[1] H. Lee Moffitt Cancer Center
[2] University of Texas, MD Anderson Cancer Center
[3] University of Minnesota
[4] Dana Farber Cancer Institute
[5] The Emmes Corporation
[6] Mount Sinai Medical Center
[7] Center for International Blood and Marrow Transplant Research, Medical College of Wisconsin
[8] University of Michigan
[9] National Heart, Lung, and Blood Institute
[10] Vanderbilt University Medical Center
[11] Fred Hutchinson Cancer Research Center
[12] Johns Hopkins University
[13] National Marrow Donor Program

Approval Signature (Protocol Chair)

Approval Signature (Protocol Officer)

Core Study Participants:

BMT Program at Northside Hospital
Baylor College of Medicine Consortium
Case Western Reserve University Consortium
City of Hope National
Dana Farber Consortium
Duke University Medical Center
Fred Hutchinson Cancer Research Center
H. Lee Moffitt Cancer Center
Johns Hopkins University
Memorial Sloan-Kettering Cancer Center
Ohio State Consortium
Pediatric BMT Consortium
University of Florida Consortium
University of Michigan Consortium
University of Minnesota
University of Nebraska Consortium
University of Pennsylvania Cancer Center
University of Texas, MD Anderson Cancer Center
Virginia Commonwealth University MCV Hospitals
Washington University, Barnes Jewish Hospital

Affiliate Study Participants:

Medical College of Wisconsin

PROTOCOL SYNOPSIS – BMT CTN PROTOCOL 1501

A Randomized, Phase II, Multicenter, Open Label Study Evaluating Sirolimus and Prednisone in Patients with Refined Minnesota Standard Risk, Ann Arbor 1/2 Confirmed Acute Graft-Versus-Host Disease

Co-Chairs: Joseph Pidala, MD, PhD and Margaret MacMillan, MD

Study Design: The study is a Phase II randomized, open label, multicenter trial designed to identify whether sirolimus is a potential alternative to prednisone as an up-front treatment for patients with standard-risk acute GVHD defined according to clinical and biomarker-based risk stratification.

Patients with previously untreated, standard-risk acute GVHD, according to the refined Minnesota Criteria, who are in need of systemic therapy, will have a 5 mL blood sample collected prior to randomization to assess their biomarker Ann Arbor Risk status. Ann Arbor scoring results will be provided 48-72 hours after randomization. Patients will begin their study treatment assignments within 24 hours of randomization. Those with biomarker results of combined AA1/2 risk will continue on their randomized study treatment and will be included for primary endpoint analysis and all planned study procedures and assessments. In contrast, patients with AA3 biomarker risk and those patients with missing biomarker results may continue on their randomized therapies or start another therapy at their physicians' discretions. In addition, AA3 risk patients and those with missing results will not be considered in primary endpoint analysis, but will be included in a subset analysis.

Primary Objective: The primary objective is to assess the rate of complete remission (CR)/partial remission (PR) on day 28 post-randomization in patients with standard-risk acute GVHD defined by both clinical and AA1/2 risk status.

Secondary Objectives: Secondary objectives are to assess the following:

1. The proportion of patients with an acute GVHD response on Day 28 (CR or PR) and who are on a prednisone (or prednisone dose-equivalent corticosteroid) dose of 0.25mg/kg/day or less.
2. Proportions of CR, PR, mixed response, no response and progression among surviving patients at Day 14, 28 and 56.
3. Treatment failure (treatment failure defined as: no response, progression, administration of additional therapy for GVHD

beyond primary therapy, or mortality) at Day 14, 28, and 56.Incidence of chronic GVHD by 6 and 12 months post-randomization.
4. Incidence of systemic infections within 6 months of randomization.
5. Freedom from acute GVHD progression, chronic GVHD, malignancy relapse and mortality at 6 months and 12 months post-randomization.
6. Disease-free and overall survival at 6 and 12 months post randomization.
7. GVHD-free survival at 6 and 12 months post-randomization.
8. Non-relapse mortality at 6 and 12 months post-randomization.

Exploratory Objectives: Exploratory objectives are to assess the following:

1. Steroid-dose (measured in prednisone-equivalent) on Days 7, 14, 21, 28, 35, 42 and 56.
2. Use of topical (skin, GI) agents for acute GVHD therapy.
3. Incidence of discontinuation of immune suppression, and immune suppression discontinuation without GVHD or disease progression/recurrence by Days 56, 180, and 365 post-therapy.
4. Incidence of EBV-associated lymphoproliferative disorder or EBV reactivation requiring therapy
5. Incidence of corticosteroid- and sirolimus-associated complications (collected in all patients):
    a. Incidence of hyperglycemia (defined as a random glucose >200mg/dL or fasting glucose >126mg/dL) and use of diabetes therapy (use of insulin and/or oral medications to control and/or maintain glucose levels) at baseline, Day 28 and Day 56.
    b. Change from baseline in functional myopathy score at Day 56 and 6-months post-randomization.
        i. Hip Flexor and Quadriceps Strength via handheld dynamometer
        ii. Two Minute Walk Test
        iii. 5-time Sit-to-Stand
        iv. Adult Myopathy Assessment Tool (AMAT)
    c. Incidence of hyperlipidemia as measured by fasting lipid panel at baseline, Days 28, 56 and 180 post-randomization.
    d. Incidence of post-transplant thrombotic microangiopathy (TMA) by 6 months post-randomization.
6. Proportion of patients requiring therapy for CMV-reactivation by day 56 post-randomization.

7. Change in patient-reported outcomes from enrollment to Day 56, 6 months and 12 months.
    a. MD Anderson Symptom Inventory (MDASI)
    b. FACT-BMT
    c. MOS Short Form 36 (SF-36)
    d. PedsQL (Pediatric patients)
8. A secondary descriptive analysis will evaluate outcomes for AA3 patients.

Eligibility: Patients of all ages with newly diagnosed standard-risk acute GVHD, diagnosed according to Refined Minnesota Criteria. All allogeneic donor sources and all conditioning regimens are allowed. Biopsy confirmation of GVHD is not required unless institutional practice mandates biopsy confirmation to make a GVHD treatment decision. Patients must have an absolute neutrophil count (ANC) greater than 500/μL. Patients must be able to tolerate oral or enterically-administered medication. Patients must have 5 mL blood samples collected for Ann Arbor Scoring. No previous systemic immune suppressive therapy for acute GVHD is allowed except topical corticosteroid use. Patients receiving sirolimus within 14 days of screening will be excluded. Patients with an active or recent (within 7 days) episode of transplant associated microangiopathy are not eligible. Patients with acute GVHD after donor lymphocyte infusion are not eligible. Patients with clinical presentation resembling de novo chronic GVHD or overlap syndrome are not eligible.

Treatment Description: Patients will be randomly assigned 1:1 to sirolimus vs. prednisone at 2mg/kg/day starting dose. Sirolimus will be loaded and then kept at maintenance dosing for target therapeutic levels for minimum duration through Day 56 post-randomization Prednisone will be kept at 2mg/kg/day x 3 days, and then tapered according to individual treating clinician judgment.

Accrual Objective: 150 total patients will be enrolled and randomized 1:1 to sirolimus vs. prednisone. It is anticipated that ~20% of randomized patients will have AA3 status or missing biomarker results resulting in 120 patients for the analysis of the primary endpoint.

Accrual Period: The estimated accrual period is 2 years.

Study Duration: Patients will be followed for 12 months following initiation of therapy.

Safety Monitoring: The rate of failure of sirolimus therapy by Day 42 post-randomization, defined as the addition of a systemic immune suppressive therapy beyond prednisone among those patients originally treated with sirolimus, will be monitored using a sequential probability ratio test (SPRT) for binary data. The SPRT will contrast a 25% and 50% 42-day rate of sirolimus failure. Day 56 mortality will also be assessed for safety monitoring using a censored exponential SPRT contrasting a 10% and 25% rate of overall mortality.

TABLE OF CONTENTS

LIST OF APPENDICES

CHAPTER 1

1. BACKGROUND AND RATIONALE

1.1. Introduction

Acute graft-versus-host disease (aGVHD) is a frequent complication of allogeneic hematopoietic cell transplantation (HCT) involving activation of donor T-lymphocytes against host tissues.[1] Despite immune suppression prophylaxis, up to 50% of HCT recipients will experience aGVHD of varying severity.[2] The syndrome of aGVHD involves multiple target organs, including the skin (presenting most frequently as a maculopapular rash), intestinal tract (presenting as nausea/vomiting and/or diarrhea), and liver (presenting as cholestatic liver injury with or without transaminase elevation). The mainstay of treatment of aGVHD for over 3 decades has been high-dose corticosteroids, typically dosed at the prednisone equivalent of 1-2 mg/kg per day.[3] Corticosteroid therapy has several shortcomings, including both limited efficacy and toxicity including infection, hyperglycemia, hypertension, hyperlipidemia, and osteoporosis.

One of the challenges that remains daunting in the clinical management of allogeneic HCT recipients is knowing which patients are likely to have mild aGVHD who would possibly be spared the side effects of prolonged high doses of corticosteroids. Different approaches to risk stratify patients with aGVHD have been recently developed and refined. These strategies include models built upon initial clinical staging of aGVHD target organs (e.g., skin, intestinal tract, and liver) and blood biomarker-based approaches. Both of these strategies hold value in identifying patients who are likely to respond well (i.e., demonstrate a complete response [CR] or partial response [PR]) to corticosteroids and thus less likely to die due to complication of the transplant (i.e., experience transplant-related mortality [TRM]).[4]

1.2. Risk Stratification by Onset Organ Severity

In 1990, Weisdorf et al. identified in multivariate analyses that overall stage score (sum of each aGVHD organ stage 0-4, plus 1 point for upper GI, for a maximum score of 13) was strongly associated with likelihood of CR.[3] Based upon this initial observation that single organ involvement was more likely to achieve a CR than multi-organ involvement, the GVHD Risk Score was subsequently developed by the Minnesota group.[5] Multivariate analysis of the outcomes of 864 consecutive patients from 1990-2007 yielded the following high-risk organ stages: skin stage 4, lower gastrointestinal stage 3+, liver stage 3+, or skin stage 3 and lower gastrointestinal or liver stage 2+ GVHD.

The GVHD Risk Score has recently been refined using data from multiple centers as well as Blood and Marrow Transplant Clinical Trials Network (BMT CTN) clinical trials 0302 and 0802 [6] with a total of 1,723 patients used in modeling – the largest aGVHD cohort analyzed for their characteristics and outcomes to date. Developed using clinical grouping and recursive partitioning, this new Risk Score (http://z.umn.edu/MNAcuteGVHDRiskScore) can classify patients into high-risk (HR) or standard risk at the onset of aGVHD symptoms. In this model, 84% of patients are classified as standard risk, defined as single organ involvement (stage 1-3 skin or stage 1-2 GI) or two organ involvement (stage 1-3 skin plus stage 1 GI; or stage 1-3 skin plus stage 1-4 liver), with a day 28 CR/PR rate of 69% (Figure 1). All others are considered HR, with a day 28 CR/PR rate of 43%. This model can be used in real time at the bedside, making it practical for stratification in clinical trials.

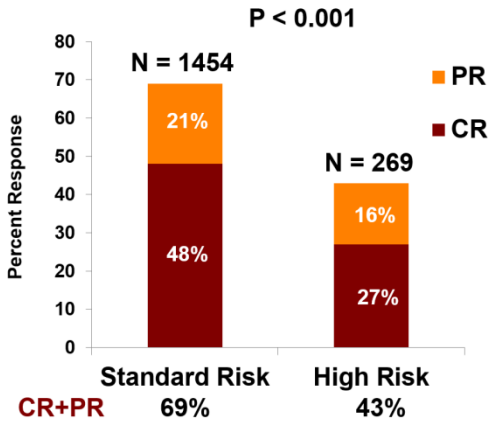


Figure 1. Comparison of day 28 CR/PR in MN Standard vs. High Risk aGVHD patients.

1.3. Risk Stratification by Blood Biomarkers

Serum proteomic patterns associated with aGVHD were first published approximately 10 years ago[7,8]. The University of Michigan group has significantly expanded the proteomic profiling of patients with aGVHD in recent years by developing a 3-level risk stratification system, Ann Arbor 1 (AA1) low-risk aGVHD, Ann Arbor 2 (AA2) intermediate-risk aGVHD, and Ann Arbor 3 (AA3) high-risk aGVHD. The Ann Arbor score is based upon serum or plasma levels of tumor necrosis factor receptor-1 (TNFR1), regenerating islet-derived 3-alpha (REG3α), and suppression of tumorigenicity 2 (ST2) measured at diagnosis of GVHD, regardless of clinical severity (grades I-IV). Each Ann Arbor score corresponds to a distinct risk of six-month non-relapse mortality (NRM), such that Ann Arbor 1 GVHD has <10% NRM, Ann Arbor 2

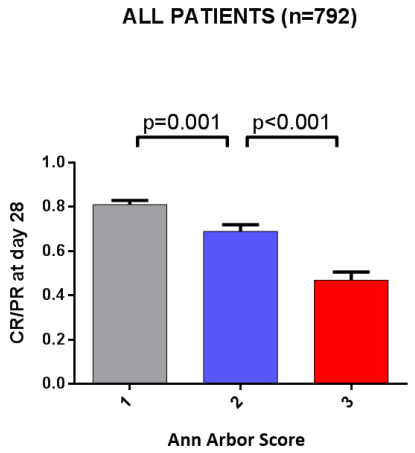
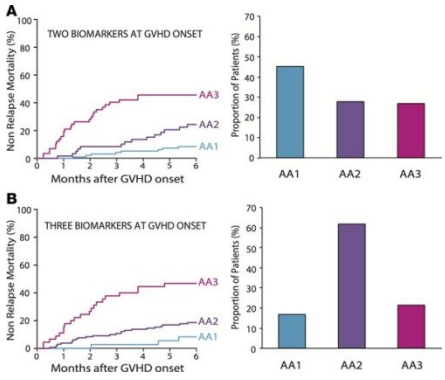


Figure 2. Comparison of day 28 CR/PR in AA1, AA2, and AA3 biomarker-based risk stratified patients.

GVHD ~25% NRM, and Ann Arbor 3 GVHD has >40% NRM. The scoring algorithm was validated in an independent test set of patients with aGVHD from the University of Michigan and the University of Regensberg, and in a separate validation set from multicenter Blood and Marrow Transplant Clinical Trials Network (BMT CTN) aGVHD treatment trials 0302 and 0802.[9] Relapse rates do not differ between Ann Arbor scores and thus the differences in non-relapse mortality translate into significant differences in survival. Differences in response to treatment account for the vast majority of differences in NRM. Patients with AA1 demonstrate 81% CR/PR at day 28 and AA2 patients demonstrate 68% CR/PR at day 28 (Figure 2). In contrast, high risk AA3 patients demonstrate 46% CR/PR at day 28. Furthermore, treatment responses for patients with AA3 GVHD are significantly less likely to be durable compared to AA1 and AA2 patients. Only 21% of AA3 patients remain in CR without flare at six months from diagnosis compared to 47% of AA1/2 patients ($p<0.001$). Importantly, similar proportions of patients are assigned to each Ann Arbor score in each of the three standard groupings of Glucksberg grades (I vs. II vs. III/IV). Thus, approximately 20% of all newly diagnosed GVHD cases have a high risk biomarker profile. These patients have poor responses to primary treatment, experience high NRM even when the clinical presentation is not severe, such as a <50% skin rash. Given their high likelihood of treatment failure, these patients are not good candidates for steroid sparing primary GVHD treatment trials. While compelling differences are observed in AA1/AA2 versus AA3 cohorts above, blood GVHD biomarkers have never previously been used to prospectively risk stratify patients.

In February 2017, the same group that developed the Ann Arbor scoring system published that the concentrations of two biomarkers, ST2 and REG3α, measured on day 7 after HCT predicted lethal GVHD. This group, now located at Mount Sinai School of Medicine, developed and validated the MAGIC (Mount Sinai Acute GVHD International Consortium) algorithm from 1287 patients who received an allogeneic HCT at 11 different centers in the United States, Europe, and Asia [Hartwell *JCI Insight* 2017]. The MAGIC two biomarker algorithm is able to predict GVHD outcomes at multiple timepoints after HCT, including day 7, day 14 (unpublished data), at GVHD onset [Hartwell, JCI Insight 2017], and after one week of GVHD treatment with systemic steroids [Major-Monfried, ASH 2016]. As shown in the figure, the MAGIC algorithm generates Ann Arbor scores with the same risk for NRM as the three biomarker algorithm, but more patients are assigned to the low risk group (Ann Arbor 1) and the high risk group (Ann Arbor 3) than with the three biomarker algorithm. As of April 7, 2017 no patient's classification as AA1/2 or AA3 would have changed if the MAGIC algorithm had been used since the activation of BMT CTN 1501. Thus, given its advantages for design of future GVHD clinical trials incorporating biomarkers and the lack of impact on this study, the MAGIC algorithm will replace the three biomarker algorithm for determining Ann Arbor risk status.



1.4. Results of Previous aGVHD Therapy Trials

To date, two multicenter aGVHD treatment trials have been conducted in the Blood and Marrow Transplant Clinical Trials Network (BMT CTN). The first, BMT CTN 0302, was a multicenter, randomized, four-arm phase II trial that was designed to identify the agent most promising for use in addition to corticosteroids for the front-line therapy of aGVHD. 180 patients with a new diagnosis of aGVHD were randomized between the addition of etanercept, mycophenolate mofetil (MMF), denileukin diftitox, or pentostatin. The proportion of day 28 CR/PR was highest in the MMF arm at 60%.[10] Thus, MMF was selected for a randomized phase III trial of MMF versus placebo for first-line therapy of aGVHD in the subsequent study, BMT CTN 0802. In CTN 0802, MMF did not meet the primary endpoint of extending GVHD-free survival at day 56 at a planned interim analysis of 235, and the trial was terminated for futility.[11] Day 28 CR/PR was approximately 50% in both arms of this study, with no statistically significant difference between the two. Neither BMT CTN 0302 nor 0802 were risk-stratified according to clinical or biomarker-based strategies.

1.5. Sirolimus as Primary, Steroid-Free aGVHD Therapy

A retrospective analysis was conducted to examine activity of sirolimus as a sole, steroid-free, acute GVHD therapy.[12] A total of 32 HCT recipients with new onset acute GVHD treated at the Moffitt Cancer Center were included. Median age was 60 (range 28-73), 72% were not in remission from their malignancy at time of HCT, and original GVHD prophylaxis consisted of tacrolimus/methotrexate (n=29), or tacrolimus/mycophenolate mofetil (n=3). Patients were treated with sirolimus as the primary therapy for acute GVHD at median of 30 (range 15 – 106) days after HCT. Acute GVHD was biopsy confirmed in 31/32 cases, and included a representative mix of target organ involvement (skin 17, 53%; GI 21, 66%, and liver 5, 16%). Overall grade was I (n=4, 13%), II (n=24, 75%), or III (n=4, 13%). None had overall grade IV. When evaluated according to Minnesota risk stratification, 27 (84%) were standard risk and 5 (16%) were high risk. Sirolimus was administered orally as a median loading dose of 6mg (range 2 to 9 mg), followed by maintenance dosing to sustain the desired target therapeutic levels (5-14ng/mL). Tacrolimus was decreased to target a range between 3-7 ng/ml. This approach was safe, and thrombotic microangiopathy (TMA) was infrequent (n=3 cases total). Per BMT CTN consensus criteria, this was grade 1 in one case, and grade 2 in two cases. TMA resolved with dose reduction of tacrolimus in all cases.

Sixteen (50%) patients achieved CR of acute GVHD (defined as sustained complete resolution GVHD for 4 weeks without addition of prednisone or other systemic immune suppressive agents) following primary therapy with sirolimus. Among these 16 patients who achieved complete resolution of acute GVHD with sirolimus alone, initial overall response (composite of partial and complete response) was achieved at median of 7 days (range 5 – 21 days) and complete response was achieved by a median of 14 days (range 5 – 28 days). In two of these cases, recurrent acute GVHD developed 7 – 12 weeks after initial CR; resolution of recurrent acute GVHD was achieved in both with the addition of 0.2 – 0.5 mg/kg body weight of prednisone.

In the remaining 16 cases, systemic glucocorticoids were initiated at a median of 9 days (range 2 – 28 days) after initiation of sirolimus with a prednisone-equivalent median dose of 0.5 (range 0.2 – 1) mg/kg, and 12 achieved resolution of acute GVHD. Uniform criteria for initiation of

systemic steroids after first line sirolimus were not employed. Prednisone was started for persistent acute GVHD manifestations of unchanged severity in 6 cases (median 9 days from sirolimus initiation, range 2-19 days), grade progression in 6 cases (median 9 days from sirolimus initiation, range 2-16 days), in the setting of partial response in 2 cases (6-7 days after sirolimus initiation), and in 2 cases for recurrent acute GVHD within 4 weeks after initial complete response to sirolimus. Four (12%) had persistent acute GVHD that was treated with mycophenolate mofetil. Of these four patients, one died following primary disease relapse, two died from non-relapse causes (sepsis, and refractory acute GVHD with sepsis), and one survived.

A matched cohort treated with standard (1mg/kg/day starting dose) prednisone for acute GVHD was assembled. These were matched to the primary sirolimus therapy cohort according to acute GVHD organ involvement and severity, as well as HLA matching, donor, graft source, and initial GVHD prophylaxis. Nineteen of these 32 subjects (59%) achieved complete remission, as compared to 16/32 (50%) of those treated with sirolimus primary therapy (p=0.47). In total, these data support safety and activity of sirolimus as a steroid-free primary therapy for acute GVHD.

1.6. Study Rationale:

Advances are needed in primary therapy of acute GVHD, as standard corticosteroid therapy is incompletely successful and associated with toxicity. Additionally, a growing body of evidence suggests that standard-risk GVHD patients can be identified using clinical and biomarker-based risk stratification tools. Lower-intensity therapy may be effective and spare toxicity in this setting. Existing data support safety and activity of sirolimus as a sole, steroid-free primary therapy in acute GVHD, however a prospective randomized trial is needed. Using a combined risk stratification approach, we will select a standard-risk (refined Minnesota standard risk, Ann Arbor 1/2 risk) acute GVHD population and conduct a randomized phase II trial examining acute GVHD response rates after primary therapy with either sirolimus or standard prednisone (2mg/kg/day starting dose, followed by taper) therapy.

CHAPTER 2

2. STUDY DESIGN

2.1. Study Overview

The study is a multicenter, Phase II, randomized trial assessing sirolimus as primary therapy for standard-risk (refined Minnesota standard risk, Ann Arbor (AA) 1/2 biomarker risk) acute GVHD. This trial incorporates both a novel up front GVHD therapy (sirolimus) as well as a novel BMT CTN developed acute GVHD biomarker. Given the two novel aspects of this trial, the purpose of the study is to estimate differences in response rates between for acute GVHD patients identified as standard risk by both clinical and biomarker risk assessment and randomized to sirolimus or prednisone. Secondary objectives will assess safety, measures of steroid burden and toxicity as well as quality of life.

2.2. Hypothesis and Specific Objectives

2.2.1. Primary Hypothesis

The primary hypothesis is that sirolimus and prednisone will achieve comparable Day 28 CR/PR rates in the treatment of standard-risk acute GVHD (defined by both refined Minnesota clinical standard-risk criteria and AA1/2 biomarker risk group).

2.3. Study Objectives

2.3.1. Primary Objective:

The primary objective is to assess the rate of complete remission (CR)/ partial remission (PR) on day 28 post-randomization in patients with standard-risk acute GVHD defined by both clinical and AA1/2 risk status..

2.3.2. Secondary Objectives are to assess the following:

1. The proportion of patients with an acute GVHD response on Day 28 (CR or PR) and who are on a prednisone (or prednisone dose-equivalent corticosteroid) dose of 0.25mg/kg/day or less.
2. Proportions of CR, PR, mixed response, no response and progression among surviving patients at Day 14, 28 and 56.
3. Treatment failure (treatment failure defined as: no response, progression, administration of additional therapy for GVHD beyond primary therapy, or mortality) at Day 14, 28, and 56.
4. Incidence of chronic GVHD by 6 and 12 months post-randomization.
5. Incidence of systemic infections within 6 months of randomization.
6. Freedom from acute GVHD progression, chronic GVHD, malignancy relapse and mortality at 6 months and 12 months post-randomization.
7. Disease-free and overall survival at 6 and 12 months post randomization.

8. GVHD-free survival at 6 and 12 months post-randomization.
9. Non-relapse mortality at 6 and 12 months post-randomization.

2.3.3. Exploratory Objectives are to assess the following:

1. Steroid-dose (measured in prednisone-equivalent) on Days 7, 14, 21, 28, 35, 42 and 56.
2. Use of topical (skin, GI) agents for acute GVHD therapy.
3. Incidence of discontinuation of immune suppression, and immune suppression discontinuation without GVHD or disease progression/recurrence by Days 56, 180, and 365 post-therapy.
4. Incidence of EBV-associated lymphoproliferative disorder or EBV reactivation requiring therapy
5. Incidence of corticosteroid- and sirolimus-associated complications (collected in all patients):
    a. Incidence of hyperglycemia (defined as a random glucose >200mg/dL or fasting glucose >126mg/dL) and use of diabetes therapy (use of insulin and/or oral medications to control and/or maintain glucose levels) at baseline, Day 28 and Day 56.
    b. Change from baseline in functional myopathy score at Day 56 and 6-months post-randomization.
        i. Hip Flexor and Quadriceps Strength via handheld dynamometer
        ii. Two Minute Walk Test
        iii. 5-time Sit-to-Stand
        iv. Adult Myopathy Assessment Tool (AMAT)
    c. Incidence of hyperlipidemia as measured by fasting lipid panel at baseline, Days 28, 56 and 180 post-randomization.
    d. Incidence of post-transplant thrombotic microangiopathy (TMA) by 6 months post-randomization.
6. Proportion of patients requiring therapy for CMV-reactivation by day 56 post-randomization.
7. Change in patient-reported outcomes from enrollment to Day 56, 6 months and 12 months.
    a. MD Anderson Symptom Inventory (MDASI)
    b. FACT-BMT
    c. MOS Short Form 36 (SF-36)
    d. PedsQL (Pediatric patients)
8. A secondary descriptive analysis will evaluate outcomes for AA3 patients.

2.4. Patient Eligibility

2.4.1. Inclusion Criteria

Biopsy of involved organs with acute graft-versus-host disease (GVHD) is encouraged, but not required for study entry. Enrollment/randomization includes commitment to continue steroids or sirolimus as specified in the protocol, as well as the required follow-up observations. *If, according to institutional practice, the intention to treat a patient depends on histologic results,*

the patient should not be enrolled on the BMT CTN 1501 study, until the biopsy results are available. Patients can be enrolled with only a clinically established diagnosis. If an unexpected biopsy result is reported, the treating physician/center should contact the protocol chair(s) to discuss further course of action.

1. Patients with their first presentation of ***standard-risk* acute GVHD**, according to refined Minnesota Criteria. Refined Minnesota Criteria are available at https://redcap.ahc.umn.edu/surveys/?s=bNmFhseJIf. **S*tandard-risk* acute GVHD according to the refined Minnesota Risk Criteria** requires meeting one of the criteria below[6]:
    a. Single organ involvement
        i. Stage 1-3 skin
        ii. Stage 1 upper GI
        iii. Stage 1-2 lower GI
    b. Multiple organ involvement
        i. Stage 1-3 skin plus stage 1 upper GI
        ii. Stage 1-3 skin plus stage 1 lower GI
        iii. Stage 1-3 skin plus stage 1 lower GI plus stage 1 upper GI
        iv. Stage 1-3 skin plus stage 1-4 liver
        v. Stage 1 lower GI plus stage 1 upper GI.
2. Acute Minnesota Standard Risk GVHD requiring systemic immune suppressive therapy.
3. Acute GVHD developing after allogeneic hematopoietic cell transplantation using either bone marrow, peripheral blood, or umbilical cord blood. Recipients of non-myeloablative, reduced intensity conditioning and myeloablative transplants are eligible. All allogeneic donor sources are permitted, including siblings, unrelated donors, HLA-haploidentical related donors and umbilical cord blood.
4. Patients that have never received systemic immune suppressive therapy for treatment of active GVHD (topical skin and GI corticosteroids are allowed).
5. Ability to tolerate oral or enterically-administered medications.
6. Patients of all ages.
7. Absolute neutrophil count (ANC) greater than 500/µL.
8. Biopsy confirmation of GVHD is not required. Enrollment should not be delayed for biopsy or pathology results unless local institutional practice mandates biopsy confirmation to make a GVHD treatment decision.
9. Written informed consent and/or assent from patient, parent or guardian.
10. Collection of a 5 ml blood sample (red top for serum) from the patient for Ann Arbor Scoring and ready to be shipped immediately after randomization.

2.4.2. Exclusion Criteria

1. Patients receiving sirolimus (for any indication including GVHD prophylaxis) within 14 days of screening for enrollment.
2. Relapsed, progressing or persistent malignancy requiring withdrawal of systemic immune suppression.
3. Patients with acute GVHD developing after a donor lymphocyte infusion.
4. Active or recent (within 7 days) episode of transplant associated microangiopathy.

5. Patients with uncontrolled infections will be excluded. Infections are considered controlled if appropriate therapy has been instituted and, at the time of enrollment, no signs of progression are present. Progression of infection is defined as hemodynamic instability attributable to sepsis, new symptoms, worsening physical signs or radiographic findings attributable to infection. Persisting fever without other signs or symptoms will not be interpreted as progressing infection.
6. Patients unlikely to be available for evaluation at the transplant center on Day 28 and 56 of therapy.
7. A clinical presentation resembling de novo chronic GVHD or overlap syndrome (as defined in Appendix C) developing before or present at the time of enrollment.
8. Patients receiving systemic corticosteroids for any indication within 7 days before the onset of acute GVHD, except the following: Stable replacement doses of corticosteroids for adrenal insufficiency are permitted (e.g. hydrocortisone total dose of 10-12 mg/m^2/day or prednisone 5-7.5mg daily or equivalent). Corticosteroids administered as premedication before transfusion of blood products or before intravenous medications to prevent infusion reactions are allowed.
9. Patients who are pregnant or breastfeeding.
10. Females of childbearing potential (FCBP) or a man who has sexual contact with a FCBP and is unwilling to use effective birth control for the duration of the study.
11. Patients on dialysis.
12. Patients on mechanical ventilation.
13. Patients with severe hepatic sinusoidal obstruction syndrome who in the judgment of the treating physician are not expected to have normalized bilirubin by Day 56 after enrollment.
14. Patients with a history of hypersensitivity to sirolimus or any component of the formulation.

2.5. Treatment Plan

2.5.1. Randomization

A 5 mL blood sample (red top tube) for Ann Arbor Scoring must be collected from the patient prior to enrollment/randomization in AdvantageEDC (see section 2.5.2). Upon confirmation of eligibility in AdvantageEDC, the study participant will be randomized in a 1:1 fashion to receive either sirolimus or prednisone. Study treatment (sirolimus/prednisone) should be initiated as soon as possible after randomization. A maximum of 24 hours from randomization to first dose of study medication is allowable.

2.5.2. Ann Arbor Scoring

Immediately after randomization, the 5 ml of blood (red top for serum) will be shipped priority overnight for early morning arrival at the Biomarker Laboratory of the Icahn School of Medicine at Mount Sinai for biomarker analysis. Samples can be shipped Monday to Friday, and results can be delivered Tuesday through Saturday. Once received in the laboratory, the GVHD biomarkers used to assign the Ann Arbor GVHD score will be measured by ELISA using standard technical procedures in a CLIA certified laboratory. Processing the sample, measuring, and confirming the ELISA assay results take 4.5 hours (range 4-6 hours). Once the Ann Arbor

score is confirmed, the investigator at the participating center will be notified if the patient has AA1/2 biomarker risk, AA3 biomarker risk or missing biomarker results. Notification of the patient's Ann Arbor results will occur within 72 hours of study enrollment/randomization (usually within 48 hours).

Those with AA1/2 status will remain on their randomized therapy. AA3 patients and patients with missing biomarker results can continue on their randomized therapy, or another therapy, per the discretion of the treating physician. All patients, regardless of treatment received or biomarker status, must complete all planned study assessments. Patients with AA1/2 biomarker status and clinical standard risk will be included in the analysis for the primary endpoint. In contrast, patients with AA3 status or missing biomarker status will not be considered for the primary analysis, but will be included as a secondary descriptive analysis. GVHD response and therapy used in patients with AA3 status may be used in exploratory analyses to generate hypotheses for future high risk GVHD trials.

2.5.3. Sirolimus (rapamycin, Rapamune®)

Drug Information

Description, Administration and Storage

Sirolimus is a naturally occurring compound produced by *Streptomyces hygroscopicus.* In addition to its immunosuppressive properties, sirolimus has antifungal, antiviral and antineoplastic properties.

*Patients can be treated with either Rapamune (sirolimus) or generic sirolimus, however it is strongly recommended to use one of these agents consistently throughout the therapy period.

1) Oral solution:

Sirolimus oral solution is supplied in cartons of 2 oz (60 mL fill) amber glass bottles, or foil pouches. The oral solution contains sirolimus at a concentration of 1 mg/mL and the following inactive ingredients: Phosal 50 PG® (phosphatidylcholine, propylene glycol, monodiglycerides, ethanol, soy fatty acids, and ascorbyl palmitate) and polysorbate 80. The oral solution also contains 1.5% - 2.5% ethanol. The appropriate dose of sirolimus oral solution should be measured using the provided amber colored oral syringe and is diluted in at least 2 oz (1/4 cup) of water or orange juice to improve palatability. No other liquids, including grapefruit juice, should be used for dilution. After vigorous mixing, the diluted dose should be taken immediately. Refill the container with an additional volume (recommended minimum of 4 oz (1/2 cup) of water or orange juice), stir vigorously, and drink or administer at once to assure delivery of all of the medication. Small children may not be able to consume the recommended volumes of water or orange juice suggested for dilution and may need lesser volumes.

Sirolimus oral solution provided in bottles may develop a slight haze when refrigerated. If such a haze occurs allow the product to stand at room temperature and shake gently until the haze disappears. The presence of this haze does not affect the quality of the product. Rapamune® Oral Solution bottles and pouches should be stored protected from

light and refrigerated at 2°C to 8°C (36°F to 46°F). The syringe should be discarded after one use. After dilution, the preparation should be used immediately.

2) Tablets:

Sirolimus tablets are available as white, tan or yellow-to-beige triangular-shaped tablets marked "RAPAMUNE 1 mg," "RAPAMUNE 0.5 mg" or "RAPAMUNE 2 mg" respectively, in bottles containing 100 tablets or cartons containing 10 blister cards each with 10 tablets (0.5 mg and 1 mg formulations only). Each tablet contains sirolimus and the following inactive ingredients: sucrose, lactose, polyethylene glycol 8000, calcium sulfate, microcrystalline cellulose, pharmaceutical glaze, talc, titanium dioxide, magnesium stearate, povidone, poloxamer 188, polyethylene glycol 20,000, glyceryl monooleate, carnauba wax, and other ingredients. The 0.5 mg and 2 mg tablets also contain yellow iron (ferric) oxide and brown iron (ferric) oxide. Sirolimus tablets should be stored at 20° to 25°C (68° - 77°F). Cartons should be used to protect blister cards and strips from light. Sirolimus tablets should be dispensed in a tight, light-resistant container. Sirolimus tablets should not be split or crushed.

3) Pharmacology:

The absorption of sirolimus is rapid after administration of Rapamune® Oral Solution, with a mean Tmax of 1-2 hours in different study populations. Oral bioavailability is only 14% in stable renal transplant patients due to first pass metabolism in the liver and the intestinal wall, plus counter transport in the gut lumen by P-glycoprotein. Mean bioavailability of sirolimus after administration of the tablet is about 27% higher relative to the oral solution. However, clinical equivalence has been demonstrated for the 2-mg dose. Co-administration with high fat meals leads to reduced Cmax, prolonged Tmax and increased AUC, meaning that sirolimus should be taken consistently with or without food. The distribution of sirolimus is notable for extensive partitioning into blood cells and approximately 92% is bound to human plasma proteins. Sirolimus is a substrate for CYP3A4 and P-glycoprotein, and is extensively metabolized by O-demethylation and/or hydroxylation to at least 7 major metabolites. The parent compound contributes to more than 90% of the immunosuppressive activity. The excretion of Sirolimus is 91% fecal and only 2.2% via the urine.

Whole blood sirolimus trough levels in renal transplant recipients who were administered daily doses of 2 mg or 5 mg at 4 hours after Neoral®) were 8.59 ± 4.01 ng/mL and 17.3 ± 7.4 ng/mL, respectively, as measured by LC/MS/MS. Whole blood trough levels significantly correlated with steady state AUC (r2=0.96). Six days of multiple dosing were required to achieve steady state. Alternatively, a loading dose of three times the maintenance dose will provide near steady state concentrations within one day in most patients.

The mean ± SD terminal elimination half life (T1/2) of sirolimus after multiple dosing in stable renal transplant patients was 62 ± 16 hours. The mean T1/2 increased from 79 ± 12 hours in subjects with normal hepatic function to 113 ± 41 hours in patients with impaired hepatic function. Dosage reduction is recommended for patients with mild to moderate hepatic impairment. Limited pharmacokinetic data are available in pediatric patients with chronically impaired renal function but indicates similar Tmax (0.62-1.6 h) and T1/2 (31-111h). Clearance is slower in males compared to females but dose

adjustments based on gender are not recommended. There were no differences between African Americans and non-African Americans.

Sirolimus administration

Patients randomized to sirolimus will receive a loading dose after which a trough sirolimus level will be measured. Based upon this sirolimus level, patients will either receive additional loading, or commence once daily maintenance sirolimus with levels checked routinely as described below.

The initial loading dose of sirolimus is required per protocol. Subsequent management of sirolimus dosing and adherence to intended therapeutic levels is strongly encouraged, but – due to inter-individual variation in drug metabolism and other factors – it is expected that drug levels can't be entirely kept within the desired range for all serial measures for all patients. In standard practice, frequent adjustment of sirolimus dosing is needed to achieve and maintain desired therapeutic levels. Thus, variation in actual dose of sirolimus and sirolimus drug levels over time will not be considered protocol deviations. The recommended dosing and drug levels serve as a practice guideline, not a protocol mandated procedure. These rules apply to dosing and monitoring of tacrolimus and cyclosporine on trial as well.

Sirolimus Loading dose

The loading dose of sirolimus is as follows:

Patients >12 years of age: 6 mg PO once
Patients ≤12 years of age: 5 mg/m^2/dose PO once (max 6 mg PO once)

Patients on voriconazole should reduce the loading dose by 50% - 90%, according to institutional practice.
Patients on posaconazole, isavuconazole and fluconazole should receive 50% of recommended loading dose.

Sirolimus Level after Loading Dose

A trough sirolimus level should be performed promptly after the loading dose (within 24-48 hours). Both initial and subsequent sirolimus level monitoring can be done both at centers participating in this protocol (if levels can be resulted within 48 hours) or through an external laboratory (also to result levels within 48 hours of receiving the sample). Study centers will work with the protocol team to ensure this is possible 7 days per week. **The target serum sirolimus level after the loading dose is 10-14 ng/mL. The following is a management guideline based upon the initial sirolimus level:**

Level	Management
< 5 ng/mL	Re-load with same loading dose
5- <10 ng/mL	Re-load with 50% of initial loading dose
10-14 ng/mL	Move to maintenance dosing

>14 ng/mL	Hold sirolimus and recheck level daily. Start maintenance therapy when level is < 14ng/mL

Initial planned maintenance dose (revised as needed based on observed sirolimus levels):
Patients >12 years of age: 2mg PO once daily
Patients ≤12 years of age: 2 mg/m^2/dose PO once daily (max 2mg daily)

Patients on voriconazole, posaconazole or isavuconazole should receive 10% of recommended maintenance dose (as an initial dose adjustment). Subsequent dose modification to maintain desired therapeutic levels will be tailored to each individual patient. When calculating dosing, review medication list for potential drug interactions. CYP-450 and p-glycoprotein drug interactions. See other dose modifications in sections below.

Maintenance Sirolimus Levels

Adults: Sirolimus levels should be performed twice weekly until at steady state, then weekly for the sirolimus therapy duration. For dose changes, recheck sirolimus level after 3 days of therapy at new dose.

Pediatrics: Due to variability in the half-life of sirolimus in children < 12 years old, levels should be checked Monday, Wednesday and Friday until 2 consistent levels, then weekly. For dose changes, recheck sirolimus level after 3 days of therapy at new dose.

Maintenance Sirolimus level goals:

- Target sirolimus level from outset through resolution of acute GVHD is 10-14ng/mL.

Level	Management
< 5 ng/mL	Re-load with 50% loading dose and increase maintenance sirolimus dose by 10-25%
5 - <10 ng/mL	Increase sirolimus dose by 10-25%
10-14 ng/mL	Continue on current dose and repeat in 1 week
>14 ng/mL	Hold sirolimus and recheck level daily. Resume therapy when level is < 14 ng/mL.

- Initial level >14-20 ng/mL: Resume therapy at 10-25% lower dose
- Initial level > 20 ng/mL: Resume therapy at 25-50% lower dose

- After acute GVHD is completely resolved, the target therapeutic range can be decreased to 5-10ng/mL

Level	Management
< 5 ng/mL	Re-load with 50% loading dose and increase maintenance sirolimus dose by 10-25%
5-10 ng/mL	Continue current dose and repeat in 1 week
>10 - 14 ng/mL	Decrease current dose by 10-25% and repeat in 1 week

>14 ng/mL	Hold sirolimus and recheck level daily. Resume therapy when level is < 10 ng/mL.

- Initial level 11-14 ng/mL: Resume therapy at 10-25% lower dose
- Initial level > 14 ng/mL: Resume therapy at 25-50% lower dose

Duration of Sirolimus:

All patients **should receive therapeutic sirolimus until at least day 56 after initiation**. Thereafter, the decision of when and how to taper is at the discretion of the treating physician.

Suggested Sirolimus Taper:

The following is a suggested sirolimus taper after completion of the required period of therapy: Reduce dose by 1/3 dose per month to complete taper off over 3 months.

Sirolimus therapy guidance:

- Sirolimus tablets should not be crushed, split or chewed.
- It is recommended to not administer oral suspension via NG or NJ, as this can obstruct the tube. Administration by G or J tube is acceptable.
- To minimize the variability of exposure to sirolimus, it should be taken at a consistent time of day and consistently with or without food.
- Fatty foods increase overall absorption (AUC).
- Sirolimus solution (1 mg/ml) may be mixed with ¼ cup of water or orange juice (no other liquids should be used). Cup should than be refilled with another 1/2 cup of water of orange juice, stirred vigorously, and then administered.
- Sirolimus solution may also be inserted in a gelatin capsule using a tuberculin syringe.
- ***Sirolimus should not be taken within 4 hours after administration of Neoral®*** *(or Gengraf®)* ***cyclosporine oral solution and/or cyclosporine gelatin capsules*** because premarketing studies in healthy volunteers demonstrated 116%-512% elevations of mean Cmax and 148%-230% elevations in AUC when sirolimus oral solution or tablets were simultaneously administered with Neoral but not when the doses were separated by 4 hours.
- ***Grapefruit juice*** reduces CYP3A4-mediated metabolism of sirolimus and must not be administered with sirolimus or used for dilution.
- ***Extreme caution when subject is on voriconazole therapy (see section 2.5.3.4)***
- ***Emesis***: The sirolimus dose may be repeated within 15 minutes of a vomited dose.
- ***Impaired renal function*** does not mandate dosage adjustment.
- ***Impaired hepatic function*** should prompt consideration for sirolimus maintenance doses to be reduced but no dose adjustment of the loading dose is necessary. Maintenance doses of sirolimus may be reduced by approximately one third in patients with hepatic impairment (Child-Pugh Score of ≥ 7/15 based on the sum of 1, 2 and 3 points respectively for each of: serum bilirubin in mg/dL (< 2, 2-3, > 3), serum albumin in mg/dL (> 3.5, 2.8-3.5, < 2.8), INR (< 1.7, 1.71-2.2, >2.2), hepatic encephalopathy (none, grade I-II, grade III-IV), or ascites (none, slight, moderate/refractory).

- ***Interchangeability of oral solution and tablets:*** Two-milligram Rapamune Oral Solution is clinically equivalent to 2-milligram Rapamune oral tablets; hence, are interchangeable on a milligram to milligram basis. However, it is not known if higher doses of Rapamune Oral Solution are clinically equivalent to higher doses of tablets on a milligram to milligram basis

Expected toxicity and management guidelines

TABLE 1. SIROLIMUS TOXICITIES

	Common >20%	**Occasional** 2-20%	**Rare** <2%
Immediate Within 1-2 days of receiving drug	Headache (L), hypertension (L), nausea, diarrhea, immuno-suppression (L), fever, constipation	Chest pain, insomnia, dyspepsia, vomiting, ***dyspnea***	Hypotension, asthma, increased cough, flu like syndrome, tachycardia, anorexia, ***hypersensitivity reactions*** (exfoliative dermatitis, angioedema)
Prompt Within 2-3 weeks, prior to the next course	Tremor (L), renal dysfunction, elevated creatinine/BUN, anemia, pain (abdominal, back, pain), ***hyperlipidemia***, hypercholesteremia, hypertrigylceridemia, hyperglycemia, ***peripheral edema***, weight gain, arthralgia	Elevated LFTs (with elevated sirolimus levels), stomatitis, urinary tract infections, URIs, mild ***thrombocytopenia, leukopenia***, hyper/hypokalemia (L), hypophosphatemia, rash, hives, pruritis, ***delayed wound healing or dehiscence (L),*** hypomagnesaemia (L), ***proteinuria***	Opportunistic infections, pleural and pericardial effusions, ***non-infectious pneumonitis or bronchiolitis-obliterans organizing pneumonia*** and pulmonary fibrosis, thrombosis, myalgias, ***increased risk of CNI-induced HUS/TTP/TMA (L)***
Delayed Any time later during therapy, excluding the above conditions	Acne		Chronic renal dysfunction, renal tubular necrosis, CHF, ascites, arthrosis, bone necrosis, osteoporosis
Late Any time after completion of treatment			Lymphoproliferative disorders, skin malignancies
Unknown Frequency and Timing	Sirolimus was embryo/fetotoxic in rats at dosages of approximately 0.2 to 0.5; clinical doses were adjusted for body surface area. It is not known whether sirolimus is excreted in human milk.		

(L) Toxicity may also occur later.

Toxicities will be scored as per the NCI's Common Terminology Criteria for Adverse Events (CTCAE) version 4.0, unless otherwise specified.

- ***Hyperlipidemia:*** In the Phase III studies, treatment of new-onset hypercholesterolemia with lipid-lowering agents was required in 42-52% of patients enrolled in the sirolimus arms. Concomitant administration of sirolimus and HMG-CoA reductase inhibitors (statins) and/or fibrates appeared to be well tolerated. However, statins or fibrates should only be administered with caution to patients being treated with sirolimus and a CNI, at doses which are reduced according to label recommendations. Treated patients should be monitored for the development of rhabdomyolysis. Sirolimus should be continued during

therapy for hypercholesterolemia and hypertriglyceridemia unless the lipid levels are uncontrollable with standard therapy and are deemed to be at risk to the subject, per the treating physician.

- ***Thrombotic Microangiopathy (TMA):*** Studies in adult transplant patients have shown an increase in TMA from 4.2% when patients were treated with tacrolimus or cyclosporine alone compared to 10.8% in patients treated with the tacrolimus/sirolimus combination.[13] TMA may occur both in the setting of sirolimus and/or CNI levels being above the therapeutic range or when levels are within desired ranges. Although complete renal recovery occurred in 92% of these patients, the potential seriousness of this complication requires careful monitoring and management.

  To meet the definition (BMT CTN Consensus Definition) for TMA, a patient must have <u>all of the following</u>:

  1. RBC fragmentation and ≥ 2 schistocytes per high-power field on peripheral smear
  2. Concurrent increased serum LDH above institutional baseline
  3. Concurrent renal* and/or neurologic dysfunction without other explanations
  4. Negative direct and indirect Coombs test results

  *Doubling of serum creatinine from baseline (baseline = creatinine before hydration and conditioning) or 50% decreased in creatinine clearance from baseline

  BMT CTN Consensus for TMA severity scoring:

Grade I	Evidence of RBC destruction (schistocytosis) without clinical consequences
Grade II	Evidence of RBC destruction with increased creatinine ≤ 3 x ULN
Grade III	Evidence of RBC destruction with creatinine > 3 x ULN not requiring dialysis
Grade IV	Evidence of RBC destruction with renal failure requiring dialysis, and/or encephalopathy

- ***Hematological Toxicity:*** ANC, hematocrit and platelet counts will be monitored at each visit. A thorough investigation should be made to assess for possible causes of the cytopenia (recurrent disease, infection, drug effect other than sirolimus, TMA, autoimmune hemolytic anemia/thrombocytopenia, GVHD, late engraftment of platelets in cord blood transplant, etc.). If other causes have been ruled out and cytopenias are significant (ANC<500, platelets <20K) sirolimus doses may be decreased by 50%. If the cytopenia does not resolve after two weeks, sirolimus may be held. If counts improve, restart sirolimus at 50% of dose and titer to full dose as tolerated.

Sirolimus drug interactions, co-administration of calcineurin inhibitors, and dose-modification for concurrent azole therapy

Drug interactions:

Sirolimus is a substrate for cytochrome CYP 3A4 and a P-glycoprotein (P-gp).
Drugs that may increase sirolimus blood concentrations include CYP 3A4, 5, or P-gp inhibitors:

- Calcineurin inhibitors: Simultaneous administration of cyclosporine soft gelatin capsules (Neoral®) results substantial increases in the sirolimus Cmax and AUC. **This is avoided if sirolimus is taken 4 hours after administration of cyclosporine.**
- Calcium Channel Blockers: diltiazem, nicardipine, nifedipine, verapamil and amlodipine. Sirolimus should be monitored and a dose adjustment may be necessary.
- Triazole antifungal agents: fluconazole, itraconazole, clotrimazole, posaconazole, isavuconazole voriconazole and ketoconazole. The magnitude of increases sirolimus C_{max}, t_{max}, and AUC is such that **sirolimus should be administered cautiously together with fluconazole, itraconazole, or posaconazole, and with extreme caution if administered together with voriconazole or isavuconazole**. If co-administration is unavoidable, then the dose of sirolimus should be greatly reduced at the time of initiation of the antifungal medication as recommended in Table 2 and that there should be very frequent monitoring of trough concentrations of sirolimus in whole blood. Sirolimus concentrations should be measured upon initiation, during co-administration, and at discontinuation of antifungal treatment, with sirolimus doses adjusted accordingly.
- Macrolide antibiotics: clarithromycin, erthyromycin, telithromycin, troleandomycin (but NOT azithromycin).
- Gastrointestinal prokinetics: cisapride, metoclopromide.

Drugs that may decrease sirolimus blood concentrations include CYP 3A4, 5, 7 or P-gp inducers:

- Anticonvulsants: carbamazepine, Phenobarbital, phenytoin.
- Antibiotics: Rifampin, rifanpentine.
- Herbs: St. John's Wort (*Hypericum perforatum*).

Care should be taken when other drugs or substances that are metabolized by CYP3A4 or P-glycoprotein are administered concomitantly with sirolimus. Grapefruit juice reduced CYP3A4 mediated metabolism of sirolimus and must not be used for dilution.

Co-administration of calcineurin inhibitors:

When co-administered with sirolimus, recommend target levels are as follows:

The target serum level for **tacrolimus** is **3-7 ng/mL**

The target serum level for **cyclosporine** is **120-200 ng/mL**

Dose adjustments are based upon clinical judgment of the managing physician after considering clinical toxicity, serum levels, GVHD, concomitant drug use and the rate of rise or decline of the serum level. Following represent *suggested* dose adjustments in calcineurin inhibitors, when initiating sirolimus to achieve target levels specified above.

Calcineurin inhibitor *actual* serum level prior to initiating sirolimus	**Recommended *target* serum level for calcineurin inhibitors when co-administered with sirolimus**	**Recommended dose adjustment to achieve recommended target serum level**
tacrolimus > 10 ng/mL	tacrolimus = 3-7 ng/mL	50% dose reduction
tacrolimus = 7.1-10 ng/mL	tacrolimus = 3-7 ng/mL	33% dose reduction
tacrolimus ≤ 7.0 ng/mL	tacrolimus = 3-7 ng/mL	No change
cyclosporine > 400 ng/mL	cyclosporine = 120-200 ng/mL	50% dose reduction
cyclosporine = 201-399 ng/mL	cyclosporine = 120-200 ng/mL	33% dose reduction
cyclosporine ≤ 200 ng/mL	cyclosporine = 120-200 ng/mL	No change

***When calcineurin inhibitors (cyclosporine, tacrolimus) are given with prednisone (i.e. among patients not randomized to the sirolimus arm of this trial), standard practices are recommended for dosing/monitoring of these agents**. Suggested target levels for these agents in the absence of sirolimus are cyclosporine 200-400ng/mL and tacrolimus 5-15ng/mL. Again, these are recommendations (not protocol mandated), and management is directed by the treating physician.

Dose-modification in setting of azole therapy:

Antifungal prophylaxis: Triazole antifungal medications are expected to increase serum CNI and sirolimus levels, therefore, dosages of CNIs and sirolimus, should be adjusted accordingly using the guidelines recommended in Table 2 and 3.

TABLE 2 - PRE-EMPTIVE DOSE REDUCTION OF SIROLIMUS OR CNIs WHEN AZOLES ARE INITIATED AT STEADY STATE LEVELS OF SIROLIMUS OR CNIs

Antifungal	**Cyclosporine**	**Tacrolimus**	**Sirolimus**
	Dose ↓	Dose ↓	Dose ↓
Voriconazole	50%	67%	90%
Posaconazole	25%	67%	90%
Isavuconazole	50%	67%	90%
Itraconazole	50%	50%	ND
Fluconazole	25%	25%	50%

***Notes:**

1. If voriconazole must be added to control fungal infection then following the 90% reduction in sirolimus dosing[14] and/or 50%-67% reduction in CNI dosing SRL and/or CNI serum levels should be measured 24-48 hours later and then every 3-4 days until levels are stable and in the desired range. If voriconazole is given intravenously or if voriconazole and sirolimus are not given together, these guidelines may not apply because the effect on bioavailability of sirolimus will be weaker.

2. Note that sirolimus tablets should not be split or crushed. Fractional dose of sirolimus may be achieved by drawing an appropriate volume of the 1 mg/mL oral liquid formulation into a 1 mL syringe and swallowed directly (or mixed with water or orange juice; no other liquids, including grapefruit juice, should be used for dilution).
3. If posaconazole or itraconazole or high-dose fluconazole are added then SRL and/or CNI serum levels should be followed 48-72 hours later and then every 3-5 days until levels are stable and in the desired range.

TABLE 3 - ANTICIPATE DOSE INCREASE OF SIROLIMUS OR CNIs WHEN AZOLES ARE STOPPED DURING CONCOMITANT SIROLIMUS OR CNI THERAPY

Antifungal	Cyclosporine		Tacrolimus		Sirolimus	
	Dose ↑	Comment	Dose ↑	Comment	Dose ↑	Comment
Voriconazole	2-fold	Dose increase may not be necessary for 5-10 days	3-fold	Dose increase may not be necessary for 5-10 days	10-fold	Dose increase may not be necessary for 5-10 days
Posaconazole	1.3-fold		3-fold		ND	
Isavuconazole	2-fold		3-fold		10-fold	
Itraconazole	2-fold		2-fold		ND	
Fluconazole	1.3-fold		1.3-fold		2-fold	

Note: Although sirolimus and CNI doses may need to be substantially increased when azole therapy is stopped, the azole mediated inhibition of cytochrome CYP 3A4 (and other) and P-glycoprotein may take 5-10 days to abate and therefore immediate dose increases are not advised. Rather, sirolimus and CNI dose increases should be cautious and based on more frequent monitoring of the sirolimus and/or CNI levels as appropriate.

Additional considerations for calcineurin inhibitor (CNI) therapy:

- ***Children*** generally require a larger total daily dose of CNI by weight than adults. If aged < 6 years, every 8 hour administration may be required to maintain desirable serum trough levels. For those patients in whom adequate serum trough levels cannot be maintained using intermittent oral or IV dosing, continuous IV infusion may be warranted.
- ***Review of concomitant medications for potential interactions that may significantly alter serum CNIs levels is essential*** because CNIs undergo extensive metabolism by the hepatic and intestinal cytochrome P-450 system which may impact toxicity and efficacy of CNIs.
- Although elevated CNI blood levels are more frequently associated with toxicity (especially renal or hepatic), ***"therapeutic" CNI blood levels may also be associated with toxicity*** (e.g. significant tremors). Therefore, dose decreases are recommended for significant organ toxicities that manifest despite a "therapeutic" CNI level.
- ***Emesis***: A repeat dose of CNI may be given if emesis happens within 15 minutes of taking the planned dose.
- CNIs should be administered at a consistent time each day and in relation to meals. Neoral oral solution should be administered in orange or apple juice at room temperature.

- Subjects should ***avoid beverages containing the enzyme bergamottin*** (grapefruit juice, Sunny Delight, Fresca, and Squirt) when taking CNIs.
- ***Distal paraesthetic pain or burning during the infusion*** may be alleviated by extending the IV infusion time for Sandimmune from the standard of 1 hour to as long as 6 hours. Alternatively, cyclosporine may be administered as a continuous infusion over 24 hours for patients intolerant of intermittent administration.
- Monitor closely for an acute allergic reaction for the first 30 minutes after starting tacrolimus IV infusion and at frequent intervals thereafter.
- ***Intravenous route of administration is preferred for patients with progressive GVHD of the liver or gastro-intestinal tract*** until GVHD is clinically improved and the oral route reliable.
- If the oral route becomes temporarily unfeasible and conversion to the intravenous formulation is required the following recommendations are made:

Starting Oral Formulation	Conversion to IV formulation
Prograf®	IV Prograf = 0.25 x total daily dose of oral Prograf but should be given as continuous daily infusion
Neoral®	IV Sandimmune = 0.4 x total daily dose of Neoral
Gengraf®	IV Sandimmune = 0.4 x total daily dose of Gengraf
Sandimmune®	IV Sandimmune = 0.25 x total daily dose of oral Sandimmune

2.5.4. Prednisone

Patients randomized to steroids will receive prednisone 2mg/kg/day PO (or methylprednisolone 1.6 mg/kg/day IV). Patients unable to take tablets can use oral prednisolone solution 2 mg/kg/day. For patients that weigh over 100kg, maximal starting dose of prednisone will be 200mg (or 2mg/kg based on a modified starting weight of 100kg). For calculation of subsequent prednisone doses/kg on subsequent measures, the modified starting weight of 100kg will be used.

Duration of Steroids and Taper

All patients should receive prednisone 2mg/kg/day PO (or methylprednisolone 1.6 mg/kg/day IV) for at least 3 full days (72 hours). This starting dose was carefully selected for the following reasons: (1) Both the clinical and biomarker-based risk stratification systems have been developed in the setting of patients treated with this starting prednisone dose; (2) The 2mg/kg/day starting dose was used in both BMT CTN 0302 and CTN 0802 trials, and expected prednisone response rates and allied power calculations for this trial are derived from this preliminary data; (3) Variable starting doses would confound the analysis of prednisone taper in this trial; (4) Prior randomized clinical trial data supported greater requirement for secondary systemic immune suppressive therapy in those treated with 1mg/kg vs. 2mg/kg starting dose.[15] Thereafter (after 72 hours of 2mg/kg/day dose), the decision of when and how to taper is at the discretion of the treating physician. A potential prednisone taper is provided in Appendix E.

2.5.5. GVHD Prophylaxis Medications

Medications such as cyclosporine and tacrolimus (if used as GVHD prophylaxis when acute GVHD developed) should be continued at therapeutic doses adjusted as necessary for renal, central nervous system (CNS) or other toxicity using conventional management guidelines.

2.5.6. Topical and Ancillary GVHD Therapies

Topical skin therapy for acute GVHD, including skin creams, as well as GI topical agents (e.g., beclomethasone or budesonide), are not encouraged on trial, however are not prohibited. Use of topical agents for management of acute GVHD will be recorded as a secondary endpoint assessment in this trial. No rules or guideline for taper and discontinuation for topical agents is provided in this protocol.

Ancillary/supportive care measures for acute GVHD such as the use of anti-motility agents for diarrhea, including octreotide, is allowed at the discretion of the treating physician. Use of ursodiol to prevent/reduce gall bladder sludging, or prevent hepatic transplant complications is also allowed according to institutional guidelines.

2.5.7. Management of study therapy after GVHD Progression or Non-response:

The following guidelines describe intended management of initial randomized therapy and thresholds for starting secondary therapy:

1. Initial randomized therapy should be given for at least 4 days before considering initiation of secondary therapy. As progressive manifestations of GVHD (new organ involvement or increased organ specific symptoms sufficient to increase the organ stage by one or more) can be ultimately stabilized and controlled in this initial therapy timeframe, secondary therapy should only be considered if GVHD continues to progress after this period.

2. If no response (no improvement in any affected organ) is observed after 7-10 days of initial randomized therapy, secondary therapy can be considered.

CHAPTER 3

3. STUDY ENDPOINTS

3.1. Primary Endpoint

The primary endpoint is the rate of complete remission (CR)/partial remissions (PR) on Day 28 post-randomization in patients with standard-risk acute GVHD (defined by both refined Minnesota clinical standard-risk criteria and AA1/2 biomarker risk group). Scoring of CR/PR are in comparison to the participant's acute GVHD status (score) on the day of randomization.

3.1.1. Response Definitions

Scoring of CR, PR, MR, NR and progression are in comparison to the participant's acute GVHD status (score) on the day of randomization.

CR is defined as a score of 0 for the GVHD grading in all evaluable organs. For example, for a response to be scored as CR at Day 56 or later, the participant must still be in CR on that day and have had no intervening additional therapy for an earlier progression, PR or NR.

Partial response (PR) is defined as improvement in one or more organs involved with GVHD symptoms without progression in others. For example, for a response to be scored as PR at Day 28 or later, the participant must still be in PR on that day and have had no intervening additional therapy for an earlier progression, PR or no response (NR).

Mixed response (MR) is defined as improvement in one or more organs with deterioration in another organ manifesting symptoms of GVHD or development of symptoms of GVHD in a new organ.

Progression is defined as deterioration in at least one organ without any improvement in others.

No response (NR) is defined as absence of any improvement or progression as defined. Patients receiving secondary therapy (including need to re-escalate steroid dose to ≥ 2.5 mg/kg/day of prednisone [or methylprednisolone equivalent of 2 mg/kg/day]) will be classified as non-responders. Patients who are assigned to the sirolimus arm who require initiation of systemic corticosteroids for the treatment of acute GVHD will be considered as non-responders for purposes of primary endpoint assessment.

3.2. Secondary Endpoints

3.2.1. Proportion of Patients with CR/PR and Steroid Dose 0.25 mg/kg or less

The proportion of patients with an acute GVHD response on Day 28 (CR or PR) and who are on a prednisone (or prednisone dose-equivalent corticosteroid) dose of 0.25mg/kg/day or less. Note

that use of prednisone below the threshold on the sirolimus arm will not be considered a failure for this endpoint.

3.2.2. Proportion of response

Proportions of CR, PR, mixed response, no response and progression among surviving patients at Day 14, 28 and 56.

3.2.3. Treatment Failure

Treatment failure is defined as: death, no response, progression, administration of additional therapy beyond primary therapy for GVHD ((or re-escalation of steroid dose to ≥ 2.5 mg/kg/day of prednisone (or methylprednisolone equivalent of 2 mg/kg/day) or initiation of corticosteroids for the treatment of acute GVHD for patients assigned to the sirolimus arm), at Day 14, 28, and 56.

3.2.4. Incidence of Chronic GVHD

Chronic GVHD is defined per NIH Consensus Criteria (see Appendix C). The incidence of chronic GVHD at 6 and 12 months post-randomization will be computed for each treatment arm, including organ involvement and severity, and overall global composite score (mild/moderate/severe).

3.2.5. Incidence of Systemic Infections

All microbiologically documented infections or significant infections requiring antibiotic/antifungal therapy occurring after initiation of therapy will be reported by site of disease, date of onset, and severity. For definitions see the BMT CTN MOP.

3.2.6. Event-Free Survival

Event-free survival is defined here as freedom from acute GVHD progression, chronic GVHD, malignancy relapse and mortality at 6 months and 12 months post-randomization.

Malignancy relapse is defined as follows:

Relapse is defined by either morphological or cytogenetic evidence of acute leukemia or MDS consistent with pre-transplant features, or radiologic evidence of lymphoma, documented or not by biopsy. Progression of disease applies to patients with lymphoproliferative diseases (lymphoma or chronic lymphocytic leukemia) not in remission prior to transplantation. The event is defined as increase in size of prior sites of disease or evidence of new sites of disease, documented or not by biopsy.

Acute leukemia and MDS – Relapse will be diagnosed when there is:

- Reappearance of leukemia blast cells in the peripheral blood; or,
- >5% blasts in the bone marrow, not attributable to another cause (e.g. bone marrow regeneration)
- The appearance of previous or new dysplastic changes (MDS specific) within the bone marrow with or without falling donor chimerism; or
- The development of extramedullary leukemia or leukemic cells in the cerebral spinal fluid or
- The reappearance of cytogenetic abnormalities present prior to transplantation

Lymphoproliferative Diseases – Relapse or progression will be diagnosed when there is:

- Appearance of any new lesion more than 1.5 cm in any axis during or at the end of therapy, even if other lesions are decreasing in size. Increased FDG uptake in a previously unaffected site will only be considered relapsed or progressive disease after confirmation with other modalities. In patients with no prior history of pulmonary lymphoma, new lung nodules identified by CT are mostly benign. Thus, a therapeutic decision should not be made solely on the basis of the PET without histologic confirmation.
- At least a 50% increase from nadir in the sum of the product diameters of any previously involved nodes, or in a single involved node, or the size of other lesions (e.g., splenic or hepatic nodules). To be considered progressive disease, a lymph node with a diameter of the short axis of less than 1.0 cm must increase by ≥ 50% and to a size of 1.5 x 1.5 cm or more than 1.5 cm in the long axis.

  Lesions should be PET positive if observed in a typical FDG-avid lymphoma or the lesion was PET positive before therapy unless the lesion is too small to be detected with current PET systems (<1.5 cm in its long axis by CT).
- In addition to the criteria above, patients with CLL who present in complete remission prior to transplantation may fulfill the relapse definition if there is reappearance of circulating malignant cells that are phenotypically characteristic of CLL.

***Institution of any therapy to treat persistent, progressive or relapsed malignancy, including the withdrawal of immunosuppressive therapy or donor lymphocyte infusion, will be considered evidence of relapse/progression regardless of whether the criteria described above were met.**

3.2.7. Disease-free and Overall Survival

Disease-free survival at 6 and 12 months post randomization. The events for disease-free survival are death and relapse of the underlying malignancy (see above – section 3.2.6 - definition for relapse/progression). Overall survival will be estimated at 6 and 12 months post-randomization.

3.2.8. GVHD-Free Survival

GVHD-free survival will be estimated at 6 and 12 months post-randomization. Both acute and chronic GVHD will be considered in this estimate.

3.2.9. Non-relapse Mortality

Non-relapse mortality at 6 and 12 months. The events for non-relapse mortality are death due to any cause other than relapse of the underlying malignancy.

3.3. Exploratory Endpoints

3.3.1. Steroid Dose

Doses of methylprednisolone will be converted to prednisone equivalents by multiplying the methylprednisolone dose by 1.25. The prednisone dose for each patient at Days 7, 14, 21, 28, 35, 42, 49, and 56 will be recorded. Prednisone doses for each patient will be converted to mg/kg. For patients that weigh over 100kg, maximal starting dose of prednisone will be 200mg (or 2mg/kg based on a modified starting weight of 100kg). For calculation of subsequent prednisone doses/kg on subsequent measures, the modified starting weight of 100kg will be used. The cumulative prednisone dose for each patient at Day 56 will be calculated by adding the doses (end of each week's dose) for each of the first eight weeks of treatment, divided by the number of days of survival during this interval.

3.3.2. Topical Therapy

The proportion of patients using either topical skin or topical GI steroids will be calculated as a secondary endpoint. Use of topical agents does not constitute systemic therapy, and does not result in designation of treatment failure.

3.3.3. Discontinuation of Immune Suppression

The dates of discontinuation of corticosteroids and other systemic immune suppressive medications will be recorded. Incidence of discontinuation of immune suppression will be assessed at Days 56, 180, and 365 post-therapy. A composite endpoint of complete immune suppression discontinuation together with freedom from any GVHD or malignancy progression/recurrence will also be examined.

3.3.4. Incidence of EBV-associated lymphoproliferative disorder

Incidence of EBV-associated lymphoproliferative disorder and EBV reactivation requiring therapy, such as reduction or withdrawal of immunosuppression, Rituximab or chemotherapy.

3.3.5. Steroid- and sirolimus- associated complications (<u>collected in all patients</u>):

Incidence of hyperglycemia

Incidence of hyperglycemia (defined as a random glucose >200mg/dL or fasting glucose >126mg/dL) and use of diabetes therapy (use of insulin and/or oral medications to control and/or maintain glucose levels) at baseline, Day 28 and Day 56.

Functional Myopathy

Change from baseline in functional myopathy score at Day 56 and 6-months post-randomization.

i. Hip Flexor and Quadriceps Strength via handheld dynamometer
ii. Two Minute Walk Test
iii. 5-time Sit-to-Stand
iv. Adult Myopathy Assessment Tool (AMAT)

Hyperlipidemia

Incidence of hyperlipidemia as measured by fasting lipid panel at baseline, Days 28, 56 and 180 post-randomization. The proportion of patients in each study arm with elevation outside of normal range for each of the measured components of the fasting lipid panel (e.g. total cholesterol, LDL, HDL, and triglycerides) will be compared. Lipid parameters and use of lipid-lowering medications will be collected.

Post-transplant thrombotic microangiopathy

Incidence of post-transplant thrombotic microangiopathy (TMA) by 6 months post-randomization. TMA definition is provided in section 2.5.3.3.

3.3.6. CMV-reactivation

Proportion of patients requiring new systemic treatment for an increasing CMV PCR level per institutional practice (patients receiving only standard of care viral prophylaxis will not be included in this assessent) for CMV-reactivation by Day 56 post-randomization.

3.3.7. Patient-reported Outcomes

These instruments will be completed by patients at enrollment, Day 56, 6 months, and 12 months. Only English speaking adult and pediatric patients, and Spanish speaking adult patients are eligible to participate in the Health Quality of Life (HQL) component of this trial. Patients >18 years will complete the FACT-BMT, MOS SF-36 and MDASI instruments. Patients $\geq$ 8 years through 18 years will complete the PedsQL™ Stem Cell Transplant Module. Surveys are completed by participants using self-completed instruments as a first choice. If this method of data collection is not possible, then surveys and response options may be read verbatim to participants, either in person or over the phone, to collect data. The method of survey completion, the date, and the language will be recorded in the database. **Surveys may not be completed by surrogates**.

Change in patient-reported outcomes from enrollment to Day 56 and 6 months will be calculated.

a. MD Anderson Symptom Inventory (MDASI)

The MDASI is a 19 item instrument that captures 13 symptoms (0="not present" to 10="as bad as you can imagine") and 6 items measuring interference with life from 0 ("did not interfere") to 10 ("interfered completely")[16]. It provides two summary scales: symptoms and interference. The MDASI will be scored according to the recommendations of the developers. We estimate it will take 5 minutes to complete the MDASI. Surveys are completed by participants using self-completed instruments as a first choice. If this method of data collection is not possible, then surveys and response options may be read verbatim to participants, either in person or over the phone, to collect data. Surveys may not be completed by surrogates.

b. FACT-BMT

The FACT-BMT, Version 4 is a self-administered instrument designed to assess multi-dimensional aspects of the QOL in BMT patients. It consists of the 27-item FACT-General (FACT-G) and the 23-item Bone Marrow Transplantation Subscale (BMTS).[17] The FACT-G assesses four primary dimensions of QOL, including physical well-being (7 items), social/family well-being (7 items), emotional well-being (6 items), and functional well-being (7 items). A five point Likert-type response scale ranging from 0 to 4 is used (0 = 'not at all'; 1 = 'a little bit'; 2 = 'somewhat'; 3 = 'quite a bit'; and 4 = 'very much'). The original FACT-BMT was developed in English using a standardized approach for item derivation, reduction and testing, and has been used extensively in various clinical trials. The FACT-BMT has since been translated into other languages.

c. MOS Short Form 36 (SF-36)

The MOS SF-36 Version 2 is a 36-item self-report questionnaire which assesses patient-reported health and functioning.[18] The instrument examines the following domains of QOL: physical functioning (PF), role functioning-physical (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role functioning-emotional (RE), and mental health (MH). Two summary scales from the SF-36 include the physical component score (PCS) and the mental component score (MCS).

d. PedsQL

The PedsQL measurement model for pediatric quality of life inventory is a multi-dimensional tool for monitoring health-related quality of life (HRQOL) in children and adolescents (age 2-18).[19] Age-specific versions for younger and older children are available. Physical, emotional, social, and school functioning are assessed. The PedsQL™ Stem Cell Transplant Module is a 46 item instrument that measures health-related quality of life in children and adolescents undergoing hematopoietic stem cell transplant, and is developmentally appropriate for self-report in ages 8 through 18 years.[20]

3.3.8. AA3 Biomarker Outcomes

A secondary descriptive analysis will evaluate outcomes for AA3 patients. Patients with a missing AA result will be excluded.

CHAPTER 4

4. PATIENT ENROLLMENT AND EVALUATION

4.1. Approaching Patients, Eligibility Screening and Obtaining Consent

Patients with previously untreated, standard-risk acute GVHD, according to the refined Minnesota Criteria, who are in need of systemic therapy, will be approached as soon as possible. Patients willing to participate in the trial will sign an Institutional Review Board approved consent form. Physicians will further evaluate each patient's eligibility for randomization onto this study (see Section 2.4).

4.2. Enrollment/Randomization

Patients will be enrolled/randomized onto the study using the BMT CTN Advantage Electronic Data Capture (EDC). Patients will be randomized to Sirolimus versus Prednisone in a 1:1 ratio. The following procedures shall be followed:

1. An authorized user at the clinical center completes the initial screening by entering patient demographics and Segment A information (consent date, inclusion/exclusion criteria and confirmation that 5 mL of blood has been collected for Ann Arbor Scoring) on the Eligibility Form.
2. If the patient is eligible, a study number and random treatment assignment is generated.
3. A visit schedule based on enrollment date is displayed for printing.
4. Immediately after randomization, the 5 ml blood sample will be shipped priority overnight for early morning arrival at the Biomarker Laboratory of the Icahn School of Medicine at Mount Sinai for biomarker analysis.

If a connection is interrupted during a randomization session, the process is completely canceled and logged. A backup manual registration and randomization system will also be available to provide for short-term system failure or unavailability.

4.2.1. Treatment

Treatment should be initiated as soon as possible after randomization. A maximum of 24 hours is allowable.

As per section 2.5.2, the patient's Ann Arbor results will be provided to the treating physician within 48-72 hours of study enrollment/randomization. Patients with AA1/2 status will continue on their randomized therapy. In contrast, patients with AA3 status and patients with missing biomarker results may continue on the randomized therapy, or another therapy per individual treating physician discretion.

4.3. Study Monitoring

4.3.1. Follow-up Schedule

The Follow-up Schedule for scheduled study visits is outlined in Table 4.3. A detailed description of each of the forms and the procedures required for forms completion and submission can be found in the Data Management Handbook and User's Guide.

Follow-up Assessments: The timing of follow-up visits is based on the date of randomization. Following randomization, the Transplant Center can print a Patient Visit Schedule listing target dates for assessments. Weeks 1-8 visits are primarily for acute GVHD scoring. The subsequent visits are for follow-up reports.

Table 4.3
FOLLOW-UP SCHEDULE

Assessment Time	Target Day[1] (Days Post-Enrollment)
1 week	7 days
2 weeks	14 days
3 weeks	21 days
4 weeks	28 days
5 weeks	35 days
6 weeks	42 days
7 weeks	49 days
8 weeks	56 days
90 days	90 days
6 months	180 days
12 months	365 days

[1]Target day range = ±3 days for Day 7 (subsequent visits through Day 56 must be scheduled weekly and within ±3 days of target date). Target day range ±14 days for Day 90 and 6 months; and ±28 days for 12 months.

Criteria for Forms Submission:
All data for patients are recorded in the electronic Case Report Forms (eCRF) exclusively designed for the study. The Principal Investigator at each of the participating centers is responsible for complete, accurate and timely reporting of data.

Criteria for timeliness of submission for all study forms are detailed in the Data Management Handbook and User's Guide. Forms that are not entered into AdvantageEDC within the specified time will be considered delinquent. A missing form will continue to be requested either until the form is entered into the AdvantageEDC and integrated into the Data and Coordinating Center's (DCC) master database, or until an exception is granted and entered into the Missing Form Exception File, as detailed in the Data Management Handbook.

Corrections in the eCRF are to be conducted only by authorized personnel and may require authorization prior to implementation of corrections. However, all earlier entries are retrievable despite corrections. All corrections are recorded automatically concerning date, time point and person. Plausibility and completeness of the eCRF are verified by personnel at the Data Coordinating Center.

At all times, the Principal Investigators at the participating centers have full responsibility for ensuring accuracy and authenticity of all clinical and laboratory data entered on the eCRFs.

Reporting Patient Deaths: The Recipient Death Information must be entered into the web-based data entry system within 24 hours of knowledge of a patient's death. If the cause of death is unknown at that time, it need not be recorded at that time. However, once the cause of death is determined, the form must be updated.

Center for International Blood and Marrow Transplant Research (CIBMTR) Data Reporting: Centers participating in BMT CTN trials must register pre and post-transplant outcomes on all consecutive hematopoietic stem cell transplants done at their institution during their time of participation to the Center for International Blood and Marrow Transplant Research (CIBMTR). Registration is done using procedures and forms of the Stem Cell Transplant Outcomes Database (SCTOD). (Note: Federal legislation requires submission of these forms for all US allotransplant recipients.) Additionally, CIBMTR pre- and post- transplant Report Forms must also be submitted for all patients enrolled on this trial according to the randomization assigned to the patient at the time of initial registration with the CIBMTR. Long-term follow-up of patients on this study will continue through routine CIBMTR mechanisms.

4.4. Assessments

All assessments are considered standard-of-care unless identified below by "*."

Prior to Enrollment/Randomization

The following pre-enrollment/randomization assessments must be completed within the designated timeframe listed below.

1. Protocol-required 5 ml of blood (red top for serum) for the determination of biomarker risk status collected immediately prior to enrollment/randomization.
2. Complete acute GVHD staging and grading information including assessments of rash, diarrhea, nausea/vomiting, weight and liver function tests; Biopsy (if done) of involved tissue (standard of care, not research) within 24 hours prior to enrollment.
3. Recording of all systemic immune suppressive therapy (as appropriate: tacrolimus, cyclosporine, etc.), as well as topical agents within 24 hours prior to enrollment.
4. History and physical exam including height and weight, patient/disease/transplantation baseline variables within 7 days prior to enrollment.
5. CBC with differential, platelet count within 7 days prior to enrollment.
6. Liver function tests (bilirubin, alkaline phosphatase, AST, ALT) plus creatinine within 7 days prior to enrollment.
7. Pregnancy test per institutional practice (if applicable) within 30 days prior to enrollment.

The 5 ml blood sample for biomarker analysis must be shipped as soon as possible after randomization to the Biomarker Laboratory of the Icahn School of Medicine at Mount Sinai (see Appendix B).

Prior to Initiation of Study Therapy

Baseline assessments must be completed within 48 hours prior to initiation of study therapy unless otherwise indicated.

1. Karnofsky or Lansky Performance Status
2. Toxicity assessment
3. Baseline patient-reported outcome measures (M.D. Anderson Symptom Inventory, FACT-BMT, SF-36, or PedsQL (Pediatrics)* Only English speaking adult and pediatric patients, and Spanish speaking adult patients are eligible to participate in the Health Quality of Life (HQL) component of this trial. Patients >18 years will complete the FACT-BMT, MOS SF-36 and MDASI instruments. Patients ≥ 8 years through 18 years will complete the PedsQL™ Stem Cell Transplant Module.
4. OPTIONAL 40 mL blood sample for research laboratory studies (see Appendix B) * Patients with weight < 13 kg will have 3 mL/kg collected.

Baseline assessments required to be done as close to enrollment (day 0) as possible

1. Baseline myopathy assessments* (may be done up to 96 hours after enrollment)
    a. Hip Flexor and Quadriceps Strength via handheld dynamometer
    b. Two Minute Walk Test
    c. 5-time Sit-to-Stand
    d. Adult Myopathy Assessment Tool (AMAT)
2. Fasting lipid profile (may be done up to 96 hours after enrollment)

Post-Randomization

The following post-randomization assessments are required of all patients, regardless of therapy received.

1. Karnofsky or Lansky performance status at Days 56, 90, 180 and 365
2. Complete acute GVHD staging and grading information including assessments of rash, diarrhea, nausea/vomiting, weight and liver function tests weekly through Day 56; Day 90, 6 and 12 months
3. Chronic GVHD evaluation (if present) Day 28, 56, 90, and 6 and 12 months
4. CBC with differential, chemistry (including liver function tests) weekly through Day 56; Day 90, 6, and 12 months
5. Fasting lipid profile at Day 28, 56, and 6 months
6. Toxicity evaluation weekly through Day 56; Day 90, 6 months and 12 months
7. Recording of all systemic immune suppression, including steroid dose and drug levels (as appropriate for sirolimus, cyclosporine, tacrolimus) weekly through Day 56, Day 90, 6, and 12 months. Second-line immune suppressive therapy (therapy beyond initial

randomized therapy of either sirolimus or prednisone) should be recorded at each of these time points.

8. Use of topical (skin, GI) steroid agents weekly through Day 56; Day 90, 6 and 12 months.
9. Data on systemic infections, and EBV PTLD (or EBV requiring therapy), recorded as per the BMT CTN (Technical) MOP through 12 months.
10. EBV & CMV monitoring as per institutional practice.
11. Myopathy assessments at Day 56 and 6 months.*
    a. Hip Flexor and Quadriceps Strength via handheld dynamometer
    b. Two Minute Walk Test
    c. 5-time Sit-to-Stand
    d. Adult Myopathy Assessment Tool (AMAT)
12. Patient reported outcomes (M.D. Anderson Symptom Inventory, FACT-BMT, SF-36, or PedsQL (Pediatrics)) at Day 56, 6 months and 1 year.* Only English speaking adult and pediatric patients, and Spanish speaking adult patients are eligible to participate in the Health Quality of Life (HQL) component of this trial. Patients >18 years will complete the FACT-BMT, MOS SF-36 and MDASI instruments. Patients ≥ 8 years through 18 years will complete the PedsQL™ Stem Cell Transplant Module.
13. OPTIONAL 30-40 mL blood samples for research laboratory studies at Day 7, 28, and 56 (see Appendix B).* Patients with weight <13 kg will have ≤3 mL/kg collected at each individual time point.

Table 4.4 – REQUIRED ASSESSMENTS

	Pre-Randomize	Pre-Therapy	Days Post Randomization										
	Baseline		**7***	**14***	**21***	**28***	**35***	**42***	**49***	**56 ***	**90****	**180*****	**365*****
History and physical exam	X												
Pregnancy test (if applicable)	X												
Karnofsky/Lansky performance status		X								X	X	X	X
Acute GVHD evaluation	X		X	X	X	X	X	X	X	X	X	X	X
Chronic GVHD evaluation						X				X	X	X	X
CBC with differential, platelet count	X		X	X	X	X	X	X	X	X	X	X	X
Basic chemistry (creatinine)	X		X	X	X	X	X	X	X[-]	X	X	X	X
Liver function tests (alkaline phosphatase, bilirubin, AST, ALT)	X		X	X	X	X	X	X	X	X	X	X	X
Fasting lipid profile		X[1]				X				X		X	
Toxicity and AE evaluation		X	X	X	X	X	X	X	X	X	X	X	X
Recording of all systemic immune suppression, including steroid dose	X		X	X	X	X	X	X	X	X	X	X	X
Drug level monitoring (e.g. tacrolimus, cyclosporine)	X	X	X	X	X	X	X	X	X	X	X	X	X
Use of second-line immune suppressive therapy			X	X	X	X	X	X	X	X	X	X	X
Use of topical agents	X		X	X	X	X	X	X	X	X	X	X	X
Malignancy relapse, death			X	X	X	X	X	X	X	X	X	X	X
Systemic infections	X		X	X	X	X	X	X	X	X	X	X	X
EBV PTLD, or EBV requiring therapy			X	X	X	X	X	X	X	X	X	X	X
CMV reactivation requiring therapy			X	X	X	X	X	X	X	X			
Myopathy assessments		X[2]								X		X	
Patient-reported outcomes: MDASI, FACT-BMT, MOS SF-36, PedsQL[3]		X								X		X	X
REQUIRED 5 mL blood sample for Ann Arbor Panel Scoring collection	X[4]												
OPTIONAL Research blood samples for ancillary studies (see Appendix B)[5]		X	X			X				X			

* +/- 3 days to allow for scheduling flexibility, holidays, etc. Subsequent visits through Day 56 must be scheduled weekly.
** +/- 14 days to allow for scheduling flexibility
*** +/- 28 days to allow for scheduling flexibility
[1] The fasting lipid panel may be done up to 96 hours after enrollment.
[2] Myopathy assessments may be done up to 96 hours after enrollment.
[3] Only English speaking adult and pediatric patients, and Spanish speaking adult patients are eligible to participate in the Health Quality of Life (HQL) component of this trial. Patients >18 years will complete the FACT-BMT, MOS SF-36 and MDASI instruments. Patients ≥ 8 years through 18 years will complete the PedsQL™ Stem Cell Transplant Module.
[4] Required blood sample for Ann Arbor Scoring must be collected prior to enrollment/randomization and shipped overnight immediately afterwards.
[5] Patients with weight < 13 kg will have ≤3 mL/kg collected

4.5. Weekly GVHD Monitoring

GVHD scoring will be performed weekly for 8 weeks from study entry. Days 0, 28 and 56 (±3 days) scoring must be performed by direct observation at the Transplant Center. Evaluations at other time points may be performed by competent clinicians other than at the Transplant Center but Transplant Center is responsible for collecting all required data.

4.6. Adverse Event Reporting

Adverse Event: An adverse event (AE) is any undesirable sign, symptom or medical condition or experience that develops or worsens in severity after starting the first dose of study treatment or any procedure specified in the protocol, even if the event is not considered to be related to the study.

Expectedness: An adverse event can be Expected or Unexpected

- **Expected adverse events** are those that have been previously identified as resulting from administration of the agent. For the purposes of this study, an adverse event is considered expected when it appears in the current adverse event list, the Investigator's Brochure, the package insert or is included in the informed consent document as a potential risk.
- **Unexpected adverse events** are those that vary in nature, intensity or frequency from information in the current adverse event list, the Investigator's Brochure, the package insert, or when it is not included in the informed consent document as a potential risk.

Serious Adverse Event: A serious adverse event (SAE), as defined by 21 CFR 312.32, is any adverse event that results in one of the following outcomes, regardless of causality and expectedness:

- **Results in death**
- **Is life-threatening.** Life-threatening means that the person was at immediate risk of death from the reaction as it occurred, i.e., it does not include a reaction which hypothetically might have caused death had it occurred in a more severe form.
- **Requires or prolongs inpatient hospitalization** (i.e., the event required at least a 24-hour hospitalization or prolonged a hospitalization beyond the expected length of stay). Hospitalization admissions and/or surgical operations scheduled to occur during the study period, but planned prior to study entry are not considered SAEs if the illness or disease existed before the person was enrolled in the trial, provided that it did not deteriorate in an unexpected manner during the trial (e.g., surgery performed earlier than planned).
- **Results in persistent or significant disability/incapacity.** Disability is defined as a substantial disruption of a person's ability to conduct normal life functions.
- **Is a congenital anomaly or birth defect;** or
- **Is an important medical event** when, based upon appropriate medical judgment, it may jeopardize the participant and require medical or surgical intervention to prevent one of

the outcomes listed above. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an emergency room or at home; blood dyscrasias or convulsions that do not result in inpatient hospitalization, or the development of drug dependency or drug abuse.

Medical and scientific judgment should be exercised in deciding whether expected reporting is also appropriate in situations other than those listed above. For example, important medical events may not be immediately life threatening or result in death or hospitalization but may jeopardize the subject or may require intervention to prevent one of the outcomes listed in the definition above (e.g., suspected transmission of an infectious agent by a medicinal product is considered a Serious Adverse Event). Any event is considered a Serious Adverse Event if it is associated with clinical signs or symptoms judged by the investigator to have a significant clinical impact.

4.6.1. Required Adverse Event Reporting

Adverse event reporting will be consistent with BMT CTN procedures (BMT CTN Administrative Manual of Procedures, Chapter 6). It is BMT CTN policy that AEs must be reported even if the investigator is unsure whether a relationship exists between the adverse event and the use of study treatment. Unexpected, serious adverse events (SAEs) will be reported through an expedited AE reporting system via AdvantageEDC. Unexpected, life-threatening and fatal SAEs must be reported within 24 hours of knowledge of the event. All other unexpected SAEs must be reported within three business days of knowledge of the event. Events entered in AdvantageEDC will be reported using NCI's Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. Expected AEs will be reported using NCI's Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0 at regular intervals as defined on the Form Submission Schedule. Any expected life-threatening SAE not collected on another study form must be reported through the expedited AE reporting system via AdvantageEDC.

The Data and Safety Monitoring Board will receive summary reports of all unexpected SAEs on a semi-annual basis.

CHAPTER 5

5. STATISTICAL CORRELATIONS

5.1. Study Design

The study is designed as a Phase II randomized, open label, multicenter trial to identify whether sirolimus is a potential alternative to prednisone as an up-front treatment for patients with standard risk acute GVHD (defined by both refined Minnesota standard clinical risk and AA1/2 biomarker risk group). The hypothesis is that sirolimus will produce comparable response rates while also reducing steroid burden in this population. As this is the first trial prospectively assessing sirolimus in a multicenter setting, the trial is not designed to formally declare non-inferiority but rather to estimate the difference in response rates between the sirolimus arm and the prednisone (or standard of care) arm. Estimation of this difference along with establishing more precise estimates for a key secondary endpoint with a response definition incorporating a measure of steroid burden (requirement a patient is on less than 0.25 mg/kg at Day 28) will allow assessment of sirolimus as a potential upfront treatment for GVHD response.

Patients will be randomized to either sirolimus or prednisone. The primary endpoint is complete or partial response (relative to their GVHD status on the day of randomization) on Day 28 post-randomization with secondary endpoints assessing alternate definitions of efficacy, safety, steroid burden and quality of life. The target enrollment is 120 eligible, randomized AA1/2 patients. One hundred and fifty patients are expected to be randomized as approximately 20% of randomized patients will have AA3 status results or missing biomarker results and be removed from the primary analysis (see the Primary Endpoint Section for details).

5.1.1. Accrual

It is estimated that accrual will take 2 years.

5.1.2. Randomization

Randomization will be performed in a 1:1 ratio using random block sizes for the two arms and stratified by transplant center.

5.1.3. Primary Endpoint

The primary endpoint of the study is defined in Chapter 3. Briefly, the primary endpoint is complete or partial response of acute GVHD (as compared to status at day of randomization) on Day 28 post-randomization. Note that patients classified as AA3 per the biomarker criteria and patients with missing biomarker results will be excluded from the primary analysis (i.e. primary and secondary endpoints) population and described in a secondary analysis.

5.1.4. Primary Hypothesis

The primary hypothesis is that up front treatment with sirolimus will result in comparable CR/PR rates as up front treatment with prednisone. No formal hypothesis testing will take place. Instead

the difference in response rates between the two arms will be estimated with a 90% confidence interval.

5.1.5. Sample Size and Power Considerations

The primary endpoint is CR/PR at Day 28 post-randomization. Background data on treatment with sirolimus is available for 27 standard risk (Minnesota Criteria) patients. The overall CR/PR rate in the population was 88.9% (95% CI 77.0%, 100.0%). If patients receiving any level of prednisone are considered failures (per Chapter 3) then the observed CR/PR rate was 55.6% (95% CI 36.8%-74.3%). Considering background data from BMT CTN 0802, data were available for 95 placebo treated patients meeting the standard risk criteria at GVHD onset. The response rate in that population was 64.2% (95% CI 54.6%-73.9%). Note that PR was not formally adjudicated by the Endpoint Review Committee for BMT CTN 0802. Instead a response of Better Grade (compared to grade at diagnosis) was adjudicated and considered here as a surrogate for PR. While the point estimates differ slightly, the confidence intervals for the two populations are overlapping. Table 5.1 shows the lower bound of a 90% asymptotic confidence interval given two scenarios: 1) success rate of 60% in each arm and 2) success rate of 60% in the prednisone arm and 50% in the sirolimus arm (sirolimus response rate is 10% lower).

Table 5.1: Lower Bound of Asymptotic 90% Confidence Interval For Difference in Response Rates Between Sirolimus and Prednisone Arms Assuming Equal Rates (60%) and a 10% Lower Response Rate in the Sirolimus Arm With a 20% Inflation Due To Biomarker Results

Randomized	Analyzed	N Analyzed Per Arm	Lower Bound of 90% Confidence Interval Assuming No Difference Between Treatment Arms*	Lower Bound of 90% Confidence Interval Assuming Sirolimus Response Rate is 10% Lower*
125	100	50	-16%	-26%
150	120	60	-15%	-25%
175	140	70	-14%	-24%
200	160	80	-13%	-23%
225	180	90	-12%	-22%
250	200	100	-12%	-22%

* Difference calculated as Sirolimus response rate minus Prednisone response rate.

The number randomized assumes that 20% of randomized patients will have biomarker results indicating high risk disease (AA3) or missing biomarker results. Note that 60 analyzable patients per arm (150 randomized) will result in a lower bound of 15% assuming that the success rates in the two populations are equivalent. If sirolimus is 10% worse than the prednisone arm, the lower bound of the confidence interval will exceed 15%. Difference in Day 28 CR/PR rate across the two arms will be reported with associated 90% confidence interval, however no pre-planned difference or width of confidence interval will be used to indicate failure of either arm or serve

as a requirement for consideration of secondary endpoints (i.e. formal non-inferiority will not be declared in this trial). A key secondary endpoint is the proportion of patients in CR/PR with a steroid dose of ≤0.25 mg/kg by Day 28. Of the 95 placebo patients, 22.1% (95% CI 13.8% - 30.5%) achieved a CR/PR by Day 28 and were receiving ≤0.25 mg/kg. This estimate may not be reliable, however, as the taper for BMT CTN 0802 suggested a minimum of 0.25 mg/kg on Day 28 as such this may be an underestimate of the true rate. Using the upper end of the confidence interval, sixty patients per arm would provide 80.6% power (two-sided 5% type 1 error rate) to test for a difference of 25% between the two groups (30% Prednisone vs. 55% Sirolimus).

5.2. Interim Analysis and Stopping Guidelines

5.2.1. Interim Analysis

There will be no interim analyses for efficacy or futility as this is a small phase II study.

5.2.2. Guidelines for Safety Monitoring

Monitoring of two key safety endpoints will be conducted monthly, and if rates significantly exceed pre-set thresholds, the NHLBI will be notified in order that the DSMB can be advised. Policies and composition of the DSMB are described in the BMT CTN's Manual of Procedures. The stopping guidelines serve as trigger for consultation with the DSMB for additional review. The key safety endpoints for this study are: 1) failure of sirolimus therapy with a failure for the stopping rule defined as the addition of a systemic immune suppressive therapy beyond prednisone among those patients originally treated with sirolimus and 2) overall mortality. Both of these endpoints will be monitored among AA1/2 patients only; failure of sirolimus therapy will be monitored in the sirolimus arm only.

5.2.2.1 Day 42 Failure of Sirolimus Therapy

The rate of sirolimus failure will be monitored up to 42 days post-randomization. At least three events must be observed in order to trigger review. The expected probability of 42 day sirolimus failure is approximately 25-30% (estimate derived from primary sirolimus monotherapy data, as well as anticipated rate of secondary therapy use beyond prednisone in prior BMT CTN aGVHD therapy trials).[11,12] Each month, the null hypothesis that the 42-day sirolimus failure rate is less than or equal to 25% is tested. For this rule, a binomial sequential probability ratio test (SPRT) will be implemented.

This sequential testing procedure conserves type I error at 5% across all of the monthly examinations for the sirolimus arm. The SPRT can be represented graphically. At each monthly interim analysis, the number of evaluable patients on study is plotted against the total number of treatment failures. The continuation region of the SPRT is defined by two parallel lines. Only the upper boundary will be used for monitoring to protect against excessive 42-day sirolimus failure. If the graph crosses the upper boundary, the SPRT rejects the null hypothesis, and concludes that there are more events than predicted by the number of evaluable patients on study. Otherwise, the SPRT continues until enrollment is complete. Only failures that occur on or before the patient has been followed for 42 days are counted.

The boundaries for the binary SPRT were constructed by setting the null rate of sirolimus failure to be 25% versus an alternative of 50% for 42-day sirolimus failure. To construct the boundary, the nominal type I and type II errors were set to be 10% and 5% respectively. The upper boundary is defined by a slope of 0.36907 and an upper intercept of 2.04921. Since the upper boundary alone is being used, the actual type I error of the binomial SPRT will be less than the nominal value without a substantial increase in type II error for the alternatives of interest (i.e. sirolimus therapy failure rates higher than expected). Table 5.2.1 illustrates the number of observed events required to cross the boundary for the binomial SPRT.

TABLE 5.2.1: Safety Monitoring Guidelines for Failure of Sirolimus Therapy

Number Evaluable Patients	Number of Events On or Prior to Day 42
4-5	4
6-7	5
8-10	6
11-13	7
14-16	8
17-18	9
19-21	10
22-24	11
25-26	12
27-29	13
30-32	14
33-35	15
36-37	16
38-40	17
41-43	18
44-45	19
46-48	20
49-51	21
52-54	22
55-56	23
57-59	24
60	25

5.2.2.2 Day 56 Overall Mortality

Day 56 overall mortality is the second key safety endpoint. This endpoint will be monitored within each treatment arm using an extension of the (SPRT) for censored exponential data, as described in greater detail below and in Appendix D. In brief, unlike the binary SPRT which assesses the event rate in the number of evaluable patients, the censored exponential SPRT considers the number of events relative to the total at risk time observed on study. The SPRT is represented graphically with the number of events on the y-axis and the total at risk time on the x-axis. The upper boundary of the SPRT will be used to guard against excess mortality. This procedure assumes a censored exponential distribution for overall mortality during the first 56 days, and censors follow-up time after 56 days.

Based on background data for standard risk acute GVHD patients, the expected Day 56 overall mortality is 13.1%. As this estimate includes AA3 patients in addition to AA1/2, to be conservative, the null for the monitoring bound will be set at 10%. An SPRT contrasting 10% vs 25% Day 56 overall mortality results in decision boundaries with a common slope of 1.18 and an upper intercept of 2.29 based on nominal type 1 and type II errors of 9% and 10% respectively with the type I and II errors selected based on the simulation study and the actual operating characteristics of the SPRT shown in Table 5.2.2 that assumed uniform accrual of 60 individuals over a 24 month time period, and exponential time to failure after randomization.

TABLE 5.2.2: OPERATING CHARACTERISTICS OF SEQUENTIAL TESTING PROCEDURE FROM A SIMULATION STUDY WITH 10,000 REPLICATIONS

Day 56 Overall Mortality

True 56-Day Rate	10%	15%	20%	25%
Probability Reject Null	0.053	0.281	0.641	0.888
Mean Month Stopped	25.2	22.1	16.8	12.0
Mean # Endpoints in 56 Days	5.8	7.6	7.9	7.0
Mean # Patients Enrolled	58.2	51.6	40.4	29.6

If the true Day 56 overall mortality rate is in line with the expected rate (i.e. 10%), then the stopping guideline has a 5% probability of being triggered. If the true overall mortality rate is higher than expected (i.e. 25% by Day 56), then the stopping guideline has an 89% probability of being triggered (on average 12 months after opening when 30 patients have been enrolled).

No monitoring of secondary therapy delivered beyond prednisone will be performed in the prednisone arm, as this treatment is considered standard of care. We will prospectively monitor the rate of AA3 participants which is expected to be 20% of standard risk patients. This rate will be reported (with a 95% confidence interval) to the DSMB at each semiannual meeting to determine if the rate is higher than expected.

5.2.3. Demographic and Baseline Characteristics

Demographics and baseline characteristics will be summarized for all patients. Characteristics to be examined are: age, gender, race/ethnicity, performance status, primary disease, risk status, transplant conditioning therapy characteristics, donor type and HLA mismatch, graft source,

initial GVHD prophylaxis delivered, donor and recipient CMV status, acute GVHD organ staging and overall grade at enrollment, topical steroid therapy use. Between group comparisons will be performed for continuous variables via a t-test (or Wilcoxon Rank-Sum) and for categorical variables, via the chi-square test (or Fisher's Exact as appropriate).

5.2.4. Analysis of Primary Endpoint

The difference in Day 28 CR/PR rates between arms will be estimated along with a 90% confidence interval. Note that patients classified as high risk per biomarker results and patients with missing biomarker results will be excluded from the primary analysis. In addition to the difference in rates, within arm CR/PR rates will be estimated with accompanying confidence intervals.

5.2.5. Analysis of Secondary Endpoints

Given the large number of secondary endpoints, a significance level of 5% will be used for all comparisons (and 95% confidence intervals). This provides some control of multiplicity without overly restricting the study's ability to detect differences.

1. Proportion of Patients with CR/PR and Steroid Dose 0.25 mg/kg or less

This is a key secondary endpoint of the trial. The rates will be estimated within each arm and comparison between arms will occur using the Z test for binomial proportions (or Barnard's Exact Unconditional as appropriate).

2. Proportion of response

Proportions of complete, partial (PR), mixed response, no response and progression among surviving patients at Day 14, 28 and 56 will be compared between the treatment groups using the chi-square test (or Fisher's Exact as appropriate).

3. Treatment Failure

Proportion of primary treatment failures at Days 14, 28 and 56 will be compared using the Z test for comparing binomial proportions (or Barnard's Exact Unconditional test as appropriate).

4. Incidence of Chronic GVHD

Cumulative incidence of chronic GVHD will be estimated (treating death and relapse as a competing risk) and curves will be compared using Gray's test. Estimates and confidence intervals will be provided for 6 and 12 months post-randomization.

5. Incidence of Systemic Infections

Frequencies of infections will be tabulated by site of disease, date of onset, and severity. The time to first serious infection will be described using the cumulative incidence function with death as the competing risk, and compared between treatments using Gray's test.

6. Event-Free Survival

Event-free survival within each treatment group will be estimated using Kaplan-Meier methodology with event time being the earliest time a patient met one of the various criteria for failure or time of last follow-up for those patients not meeting the failure criteria. Pre-specified

time points of interest are 6 and 12 months post-randomization. Estimates and confidence intervals will be calculated for these time points.

7. Overall Survival

Overall survival will be estimated using Kaplan-Meier methodology with event time being the earliest time of death or last follow-up. Pre-specified time points of interest are 6 and 12 months post-randomization. Estimates and confidence intervals will be calculated for these timepoints.

8. Disease-Free Survival

Disease-free survival will be estimated using Kaplan-Meier methodology with event time being the earliest time a patient experienced a relapse of underlying malignancy, death or last follow-up. Pre-specified time points of interest are 6 and 12 months post-randomization. Estimates and confidence intervals will be calculated for these time points.

9. GVHD-Free Survival

The proportion of patients alive and GVHD free at 6 and 12 months post-randomization will be estimated and compared between treatment groups using the Z test for comparing binomial proportions (or Barnard's Exact Unconditional test as appropriate).

10. Non-relapse Mortality

Cumulative incidence of non-relapse mortality will be estimated (treating relapse as a competing risk) and curves will be compared using Gray's test. Estimates and confidence intervals will be provided for 6 and 12 months post-randomization.

5.2.6. Analysis of Exploratory Endpoints

The following endpoints and subgroup analyses are considered exploratory and will be analyzed at a significance level similar to the secondary endpoints.

1. Steroid Dose

Cumulative steroid dose will be described and compared between the treatment groups using a Wilcoxon Rank Sum test.

2. Topical Therapy

The proportion of patients using either topical skin or topical GI steroids will be estimated within treatment group and by use at randomization. Among patients entering with no history of topical therapy, the proportion of patients initiating new topical treatment will be compared between treatment groups using the Z test for comparing binomial proportions (or Barnard's Exact Unconditional test as appropriate).

3. Discontinuation of Immune Suppression

Cumulative incidence of immune suppression discontinuation will be estimated (treating death as a competing risk) and curves will be compared using Gray's test. Estimates and confidence intervals will be provided for Day 56, 6 and 12 months post-randomization.

4. Incidence of EBV-associated lymphoproliferative disorder

The incidence of EBV-associated lymphoproliferative disorder or EBV reactivation therapy will be described using the cumulative incidence function, treating death as the competing risk. Cumulative incidence curves will be compared between treatments using Gray's test.

5. Incidence of hyperglycemia

Incidence of hyperglycemia and use of diabetes therapy at baseline, Day 28 and Day 56 will be estimated and compared between arms using a Z-test (or Barnard's Exact Unconditional as appropriate).

6. Functional Myopathy

Change from baseline in functional myopathy score at Day 56 and 6-months post-randomization.

i. Hip Flexor and Quadriceps Strength via handheld dynamometer
ii. Two Minute Walk Test
iii. 5-time Sit-to-Stand
iv. Adult Myopathy Assessment Tool (AMAT)

These measures will first be assessed in patients with available data. Change from baseline within treatment group will be assessed using a Wilcoxon Signed Rank test at each timepoint. Change from baseline at each timepoint will be compared between treatment groups using a Wilcoxon Rank Sum test.

7. Hyperlipidemia

Prevalence of hyperlipidemia as measured by fasting lipid panel and use of lipid-lowering agents at baseline, Days 28, 56 and 180 post-randomization will be compared between groups at each timepoint using a Z-test for binomial proportions (or Barnard's Exact Unconditional test as appropriate).

8. Post-transplant thrombotic microangiopathy

Cumulative incidence of post-transplant thrombotic microangiopathy will be estimated (treating death as a competing risk) and curves will be compared using Gray's test. Estimates and confidence intervals will be provided for 6 months post-randomization.

9. CMV-reactivation

Proportion of patients requiring therapy for CMV-reactivation by Day 56 post-randomization will be compared between treatment groups using a Z-test (or Barnard's Exact unconditional as appropriate).

10. Change in Patient-Reported Outcomes from Enrollment to Day 56, 6 months and 1 year.

Participant self-reported measures will be assessed using the MDASI, FACT-BMT, and MOS SF-36 (or PedsQL for Pediatric patients) at enrollment, day 56, 6 months and 1 year. These will be scored according to the recommendations of the developers. Patient reported outcomes at each time point will be summarized using simple descriptive statistics (mean, SD). PRO among survivors at each time point will be compared between treatment arms in an initial analysis using two sample t-statistics. The missing data pattern of the PRO measurements will be examined using graphical techniques and logistic regression models conditional on survival. At each time point, estimates of the difference in PRO between the treatments conditional on survival at that

time point will be obtained using inverse probability of censoring weighting with independent estimating equations to account for missing data.

5.2.7. Secondary Analysis of AA3 Patients:

A secondary analysis will consider study outcomes for AA3 patients (note patients with missing biomarker results will be excluded). Roughly 30 AA3 patients are expected to be enrolled with 15 receiving sirolimus while awaiting results and 15 receiving prednisone. Outcomes for these patients will be described by initial treatment to assess whether up front sirolimus vs. prednisone impacted outcome. Due to small numbers, the analysis will be primarily descriptive as we are not likely to have power to detect differences between the groups.

5.2.8. Exploratory Subgroup Analysis:

Rates of Day 28 CR/PR (and 95% confidence intervals) will be estimated for each treatment group by donor type, graft source, and HLA mismatch (matched vs. mismatched). As the numbers for these subgroup analyses will be small, the intent is not to formally declare differences between subgroups, but rather to identify the potential for differences that may need to be accounted for in future trials.

APPENDIX A

HUMAN SUBJECTS

APPENDIX A

HUMAN SUBJECTS

Subject consent: The Network will provide templates of the consent and assent forms to each center. Each center will customize the templates according to their local requirements and submit for review by the DCC for adequacy prior to submitting to the local Internal Review Board (IRB) or the NMDP IRB of Record. Each center must provide evidence of IRB approval of the protocol and consent/assent forms.

Candidates for the study will be identified as described in Chapter 4 of the protocol. The Principal Investigator or his/her designee at each transplant center will contact the candidates, provide the patient with information about the purpose of the study and obtain consent. A trained person will enroll/randomize the patient in the AdvantageEDC system.

Confidentiality: Confidentiality will be maintained by individual names being masked and assigned a patient identifier code. The code relaying the patient's identity with the ID code will be kept separately at the center. The ID code will be transmitted to the network.

Participation of women, children, minorities and other populations: Women, children and ethnic minorities will be included in this study.

Accrual will be monitored within each center with the expectation that the enrolled patient population is representative of the transplanted patient population at each center. Representation will be examined by comparing gender, race, ethnicity and age distributions. Accrual of minority patients will be expected to be in proportion to the number of minority patients transplanted at each center. The DCC and NHLBI will discuss enrollment anomalies with the centers.

APPENDIX B

LABORATORY PROCEDURES

APPENDIX B

LABORATORY PROCEDURES

A. Blood Samples for Protocol-required GVHD Risk Assessment – Ann Arbor Biomarker Panel Scoring

Once consented, a 5 mL of blood sample (red top for serum) will be collected from the patient prior to randomization. The blood sample will be shipped on the day of collection after randomization by priority overnight FedEx for early morning arrival at the Biomarker Laboratory of the Icahn School of Medicine at Mount Sinai for biomarker panel analysis. Samples can be shipped Monday to Friday each week, and results can be delivered back to submitting clinical centers Tuesday through Saturday of each week.

Once received in the laboratory, the GVHD biomarkers used to assign the Ann Arbor GVHD score will be measured by ELISA using standard technical procedures in a CLIA certified laboratory. Processing the sample, measuring, and confirming the ELISA assay results takes 4.5 hours (range 4-6 hours). Once the Ann Arbor score is confirmed, the investigator at the participating center will be notified if the patient has Ann Arbor 1/2, Ann Arbor 3 or missing biomarker risk status *by telephone with email confirmation*. Treating physicians will be notified of the patient's Ann Arbor GVHD score within 72 hours of study enrollment and randomization (usually within 48 hours), and will receive a written laboratory report for the patient's chart detailing the final Ann Arbor score assignment based upon the expert review of testing results and algorithm output.

GVHD therapy for Ann Arbor 3 patients, which will be at the treating physician's discretion, may be used in exploratory analyses to generate hypotheses for future high risk GVHD trials.

B. Optional Research Samples Supporting Future Biomarker Studies

Optional research blood samples will be drawn at baseline (pre-treatment), and then at the serial time points outlined in the table below.

B.1 Rationale for type and schedule of planned optional research samples

1. Optional Baseline samples

Optional Baseline samples will be drawn prior to initiation of randomized therapy. While samples for subsequent RNA studies will be collected in PAXgene tubes (to stabilize RNA), other samples will be shipped to the BMT CTN Repository for processing, and aliquot storage.

The major planned use of these optional research samples will be for discovery of new RNA and protein biomarkers of therapeutic response and subsequent mortality among standard risk acute GVHD patients. Both single time point baseline values and change from baseline to Day 7 values will be considered. Studies will be performed separately among sirolimus-treated or prednisone-treated patients, and commonality of predictive markers will be explored. While protein markers

of GVHD response and subsequent mortality have been well established in the setting of standard prednisone primary therapy, this has not been addressed in sirolimus primary therapy before. Thus, this trial offers a completely new opportunity for discovery. The concurrent collection of research blood samples from AA3 patients permits comparisons across the AA1/2 vs. AA3 patients, both in the setting of sirolimus monotherapy and standard prednisone therapy.

Additional PBMC isolated for immune reconstitution research will at minimum be used for establishing baseline pre-treatment levels of regulatory T cells (Treg). Change from this baseline value to subsequent measures (including Days 7, 28, and 56) will be compared across the sirolimus- and prednisone-treated patients. We hypothesize that sirolimus therapy will support greater percentage and absolute number of Treg, as well as Treg/Tconv ratio, over the therapy period in comparison to the prednisone-treated group. As well, correlation between Treg and Treg/Tconv ratio and GVHD response and subsequent outcomes will be examined.

In addition to these studies, stored samples will be available for use by additional investigators. Submitted proposals for use of stored samples will be first reviewed by the protocol team and further adjudicated according to the procedures detailed in the BMT CTN MOP.

2. Optional Serial samples following initiation of randomized therapy

Sample types and procedures for optional serial follow up samples mirror those outlined for baseline samples.

Change from baseline to subsequent samples (e.g. Day 7) will be examined in the predictive biomarker analyses outlined above. In particular, change in serum protein biomarkers from baseline to Day 7 will be prioritized.

Single time point samples at Day 28 and 56 will also be utilized for studies examining association between RNA and protein biomarkers and Day 28 or 56 GVHD response categories outlined in the protocol. These interval time point samples (e.g. Day 28, Day 56) can also be used for prediction of later outcomes, including development of chronic GVHD, as well as development of immune tolerance as measured by freedom from GVHD and complete discontinuation of immune suppression. We anticipate that – with an anticipated median onset time of acute GVHD of approximately 20-30 days post-transplant – that the Day 56 sample (post-acute GVHD onset) will largely occur around 90 days post-transplant. This sample will be ideal for chronic GVHD predictive analyses, as this should precede median onset time for chronic GVHD (expected around 4-6 months post-HCT). Finally, these later sample time points will permit an extended view of Treg reconstitution from therapy initiation onward over a several month period. We will compare Treg reconstitution across study arms at both single time points, as well as the trajectory from baseline onward.

B.2 Optional Research Sample procedures

All optional research sample aliquots will be given unique bar code designations that cannot be linked back to the participant's name or other identifying information. Laboratory test results, clinical information, etc., associated with the coded samples may be provided to the Investigators to associate biological findings with clinical outcomes. Investigators will be able to link results

from serial serum sample aliquots (for example: Pre-treatment, Day 7, Day 28 post-treatment initiation samples). Similarly, investigators will be able to link data from different sample types. All optional research samples will be collected and shipped same-day (Monday-Friday) to the BMT CTN Biorepository for next-day processing (Tuesday-Saturday) and sample aliquot storage. Sample collection and shipping procedures are detailed in the BMT CTN 1501 Research Sample Information Guide.

OPTIONAL Research Sample Collection Schedule, Processing and Aliquot Storage Summary

Optional Research Samples (collected from all patients that provided consent for research samples)

Time Points	Sample Quantity[1]	Stored Material	Sample Processing & Storage Site	Aliquots Stored	Purpose
Pre-Treatment & Post- Initiation of Treatment Day 28 ± 3 Day 56 ± 3	10 mL PAXgene	Whole Blood Lysate	BMT CTN Biorepository	Maximum **4 aliquots** 2.5 mL-fill PAXgene tubes; stored at -80º C	Gene Expression Profile Research
Pre-Treatment & Post- Initiation of Treatment Day 7 ± 3 Day 28 ± 3 Day 56 ± 3	20 mL Heparin	Viable PBMC	BMT CTN Biorepository	Maximum **6 aliquots** 1.0 mL aliquots containing ~ 2.5-5.0 x 10^6 PBMC; controlled-rate frozen and stored in LN2	Immune Reconstitution Research
	10 mL Serum Clot Tube	Serum	BMT CTN Biorepository	Maximum **10 aliquots** ~ 0.5 mL aliquots; stored at -80º C	Proteomic and miRNA Biomarker Research

[1] **Sample Quantity** Patients with weight < 13 kg will have ≤ 3 mL/kg collected at each individual time point Research samples should be drawn in the following descending order of priority to total permissible volume: (1) 20mL heparin tube for PBMC, (2) 10mL serum clot tube for serum, and (3) 10mL PAXgene tube for whole blood lysate

APPENDIX C

DIAGNOSIS AND SEVERITY SCORING FOR ACUTE AND CHRONIC GVHD

APPENDIX C

DIAGNOSIS AND SEVERITY SCORING FOR ACUTE AND CHRONIC GVHD

1. GVHD organ staging[21]

Stage	Skin (Active Erythema Only)	Liver (Bilirubin)	Upper GI	Lower GI (stool output/day)
0	No active (erythematous) GVHD rash	<2 mg/dL	No or intermittent nausea, vomiting, or anorexia	Adult: <500 mL/day or <3 episodes/day Child: <10 mL/kg/day or <4 episodes/day
1	Maculopapular rash <25% BSA	2-3 mg/dL	Persistent nausea, vomiting or anorexia	Adult: 500-999 mL/day or 3-4 episodes/day Child: 10-19.9 mL/kg/day or 4-6 episodes/day
2	Maculopapular rash 25-50% BSA	3.1-6 mg/dL		Adult: 1000-1500 mL/day or 5-7 episodes/day Child: 20-30 mL/kg/day or 7-10 episodes/day
3	Maculopapular rash >50% BSA	6.1-15 mg/dL		Adult: >1500 mL/day or >7 episodes/day Child: >30 mL/kg/day or >10 episodes/day
4	Generalized erythroderma (>50% BSA) *plus* bullous formation and desquamation >5% BSA	>15 mg/dL		Severe abdominal pain with or without ileus or grossly bloody stool (regardless of stool volume).

Overall clinical grade (based on most severe target organ involvement):
Grade 0: No stage 1-4 of any organ.
Grade I: Stage 1-2 skin without liver, upper GI, or lower GI involvement.
Grade II: Stage 3 rash and/or stage 1 liver and/or stage 1 upper GI and/or stage 1 lower GI.
Grade III: Stage 2-3 liver and/or stage 2-3 lower GI, with stage 0-3 skin and/or stage 0-1 upper GI.
Grade IV: Stage 4 skin, liver, or lower GI involvement, with stage 0-1 upper GI.

2. Grading of Chronic GVHD (NIH Criteria)[22]

	SCORE 0	SCORE 1	SCORE 2	SCORE 3
PERFORMANCE SCORE: ☐ **KPS ECOG LPS**	☐ Asymptomatic and fully active (ECOG 0; KPS or LPS 100%)	☐ Symptomatic, fully ambulatory, restricted only in physically strenuous activity (ECOG 1, KPS or LPS 80-90%)	☐ Symptomatic, ambulatory, capabl of self-care, >50% of waking hours ou of bed (ECOG 2, KPS or LPS 60-70%)	☐ Symptomatic, limited self-care, >50% of waking hours in bed (ECOG 3-4, KPS or LPS <60%)
SKIN† **SCORE % BSA** ☐ *GVHD features to be scored by BSA:* **Check all that apply:** ☐ Maculopapular rash/erythema ☐ Lichen planus-like features ☐ Sclerotic features ☐ Papulosquamous lesions or ichthyosis ☐ Keratosis pilaris-like GVHD	☐ No BSA involved	☐ 1-18% BSA	☐ 19-50% BSA	☐ >50% BSA
SKIN FEATURES SCORE:	☐ No sclerotic features		☐ Superficial sclerotic features "not hidebound" (able to pinch)	**Check all that apply:** ☐ Deep sclerotic features ☐ "Hidebound" (unable to pinch) ☐ Impaired mobility ☐ Ulceration
Other skin GVHD features (NOT scored by BSA) **Check all that apply:** ☐ Hyperpigmentation ☐ Hypopigmentation ☐ Poikiloderma ☐ Severe or generalized pruritus ☐ Hair involvement ☐ Nail involvement ☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*				
MOUTH *Lichen planus-like features present:* ☐ **Yes** ☐ **No**	☐ No symptoms	☐ Mild symptoms **with** disease signs but not limiting oral intake significantly	☐ Moderate symptoms with disease signs **with** partial limitation of oral intake	☐ Severe symptoms with disease signs on examination **with** major limitation of oral intake
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*				

Organ scoring of chronic GVHD. ECOG indicates Eastern Cooperative Oncology Group; KPS, Karnofsky Performance Status; LPS, Lansky Performance Status; BSA, body surface area; ADL, activities of daily living; LFTs, liver function tests; AP, alkaline phosphatase; ALT, alanine aminotransferase; ULN, normal upper limit. *Weight loss within 3 months. Skin scoring should use both percentage of BSA involved by disease signs and the cutaneous features scales. When a discrepancy exists between the percentage of total body surface (BSA) score and the skin feature score, OR if superficial sclerotic features are present (Score 2), but there is impaired mobility or ulceration (Score 3), the higher level should be used for the final skin scoring. To be completed by specialist or trained medical providers. **Lung scoring should be performed using both the symptoms and FEV1 scores whenever possible. FEV1 should be used in the final lung scoring where there is discrepancy between symptoms and FEV1 scores.

	SCORE 0	SCORE 1	SCORE 2	SCORE 3
EYES *Keratoconjunctivitis sicca (KCS) confirmed by ophthalmologist:* ☐ **Yes** ☐ **No** ☐ **Not examined**	☐ No symptoms	☐ Mild dry eye symptoms not affecting ADL (requirement of lubricant eye drops ≤ 3 x per day)	☐ Moderate dry eye symptoms partially affecting ADL (requiring lubricant eye drops > 3 x per day or punctal plugs), **WITHOUT** new vision impairment due to KCS	☐ Severe dry eye symptoms significantly affecting ADL (special eyeware to relieve pain) **OR** unable to work because of ocular symptoms **OR** loss of vision due to KCS
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*				
GI Tract ***Check all that apply:*** ☐ Esophageal web/ proximal stricture or ring ☐ Dysphagia ☐ Anorexia ☐ Nausea ☐ Vomiting ☐ Diarrhea ☐ Weight loss ≥5%* ☐ Failure to thrive	☐ No symptoms	☐ Symptoms without significant weight loss* (<5%)	☐ Symptoms associated with mild to moderate weight loss* (5-15%) **OR** moderate diarrhea without significant interference with daily living	☐ Symptoms associated with significant weight loss* >15%, requires nutritional supplement for most calorie needs **OR** esophageal dilation **OR** severe diarrhea with significant interference with daily living
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____				
LIVER	☐ Normal total bilirubin and ALT or AP < 3 x ULN	☐ Normal total bilirubin with ALT ≥3 to 5 x ULN or AP ≥ 3 x ULN	☐ Elevated total bilirubin but ≤3 mg/dL or ALT > 5 ULN	☐ Elevated total bilirubin > 3 mg/dL
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*				
LUNGS** **Symptom score**:	☐ No symptoms	☐ Mild symptoms (shortness of breath after climbing one flight of steps)	☐ Moderate symptoms (shortness of breath after walking on flat ground)	☐ Severe symptoms (shortness of breath at rest; requiring 0_2)
Lung score: **%** FEV1 ☐	☐ FEV1≥80%	☐ FEV1 60-79%	☐ FEV1 40-59%	☐ FEV1 ≤39%
Pulmonary function tests ☐ Not performed				
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____				

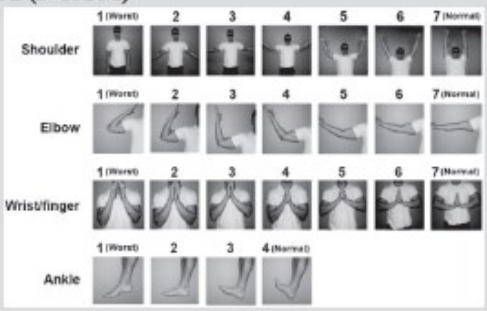
	SCORE 0	SCORE 1	SCORE 2	SCORE 3
JOINTS AND FASCIA P-ROM score *(see below)* Shoulder (1-7):___ Elbow (1-7): ____ Wrist/finger (1-7):___ Ankle (1-4):___	☐ No symptoms	☐ Mild tightness of arms or legs, normal or mild decreased range of motion (ROM) **AND** not affecting ADL	☐ Tightness of arms or legs **OR** joint contractures, erythema thought due to fasciitis, moderate decrease ROM **AND** mild to moderate limitation of ADL	☐ Contractures **WITH** significant decrease of ROM ***AND*** significant limitation of ADL (unable to tie shoes, button shirts, dress self etc.)
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____				
GENITAL TRACT (*See Supplemental figure*‡) ☐ Not examined *Currently sexually active* ☐ Yes ☐ No	☐ No signs	☐ Mild signs‡ and females with or without discomfort on exam	☐ Moderate signs‡ and may have symptoms with discomfort on exam	☐ Severe signs‡ with or without symptoms
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____				

Other indicators, clinical features or complications related to chronic GVHD (check all that apply and assign a score to severity (0-3) based on functional impact where applicable none – 0,mild -1, moderate -2, severe – 3)

☐ Ascites (serositis)___	☐ Myasthenia Gravis___	
☐ Pericardial Effusion___	☐ Peripheral Neuropathy___	☐ Eosinophilia > 500/µl___
☐ Pleural Effusion(s)___	☐ Polymyositis___	☐ Platelets <100,000/µl ___
☐ Nephrotic syndrome___	☐ Weight loss>5%* without GI symptoms___	☐ Others (specify):_____

Overall GVHD Severity *(Opinion of the evaluator)* ☐ **No GVHD** ☐ **Mild** ☐ **Moderate** ☐ **Severe**

Photographic Range of Motion (P-ROM)



3. Categories of Acute and Chronic GVHD

Categories of Acute and Chronic GVHD			
Category	**Time of Symptoms after HCT**	**Presence of Acute GVHD Features**	**Presence of Chronic GVHD Features***
Acute GVHD			
Classic acute GVHD	≤100 d	Yes	No
Late-onset acute GVHD	>100 d	Yes	No
Chronic GVHD			
Classic chronic GVHD	No time limit	No	Yes
Overlap syndrome	No time limit	Yes	Yes

*As defined in **section 4** (below)

4. Signs and Symptoms of Chronic GVHD[22]

Organ or Site	Diagnostic (Sufficient to Establish the Diagnosis of chronic GVHD)	Distinctive* (Seen in chronic GVHD, but Insufficient Alone to Establish a Diagnosis)	Other Features or Unclassified Entities†	Common‡ (Seen with Both Acute and chronic GVHD)
Skin	Poikiloderma Lichen planus–like features Sclerotic features Morphea-like features Lichen sclerosus–like features	Depigmentation Papulosquamous lesions	Sweat impairment Ichthyosis Keratosis pilaris Hypopigmentation Hyperpigmentation	Erythema Maculopapular rash Pruritus
Nails		Dystrophy Longitudinal ridging, splitting or brittle features Onycholysis Pterygium unguis Nail loss (usually symmetric, affects most nails)		
Scalp and body hair		New onset of scarring or nonscarring scalp alopecia (after recovery from chemoradiotherapy) Loss of body hair Scaling	Thinning scalp hair, typically patchy, coarse or dull (not explained by endocrine or other causes) Premature gray hair	
Mouth	Lichen planus–like changes	Xerostomia Mucoceles Mucosal atrophy Ulcers Pseudomembranes		Gingivitis Mucositis Erythema Pain
Eyes		New onset dry, gritty, or painful eyes Cicatricial conjunctivitis KCS Confluent areas of punctate keratopathy	Photophobia Periorbital hyperpigmentation Blepharitis (erythema of the eyelids with edema)	
Genitalia	Lichen planus–like features Lichen sclerosus–like features	Erosions Fissures		
Females	Vaginal scarring or clitoral/labial agglutination	Ulcers		
Males	Phimosis or urethral/meatus scarring or stenosis			
GI Tract	Esophageal web Strictures or stenosis in the upper to mid third of the esophagus		Exocrine pancreatic insufficiency	Anorexia Nausea Vomiting Diarrhea Weight loss Failure to thrive (infants and children
Liver				Total bilirubin, alkaline phosphatase > 2 × upper limit of normal ALT > 2 × upper limit of normal
Lung	Bronchiolitis obliterans diagnosed with lung biopsy BOS§	Air trapping and bronchiectasis on chest CT	Cryptogenic organizing pneumonia Restrictive lung disease‖	
Muscles, fascia, joints	Fasciitis Joint stiffness or contractures secondary to fasciitis or sclerosis	Myositis or polymyositis¶	Edema Muscle cramps Arthralgia or arthritis	
Hematopoietic and Immune			Thrombocytopenia Eosinophilia Lymphopenia Hypo- or hyper-gammaglobulinemia Autoantibodies (AIHA, ITP) Raynaud's phenomenon	
Other			Pericardial or pleural effusions Ascites Peripheral neuropathy Nephrotic syndrome Myasthenia gravis Cardiac conduction abnormality or cardiomyopathy	

ALT indicates alanine aminotransferase; AIHA, autoimmune hemolytic anemia; ITP, idiopathic thrombocytopenic purpura.
* In all cases, infection, drug effect, malignancy, or other causes must be excluded.
† Can be acknowledged as part of the chronic GVHD manifestations if diagnosis is confirmed.
‡ Common refers to shared features by both acute and chronic GVHD.
§ BOS can be diagnostic for lung chronic GVHD only if distinctive sign or symptom present in another organ (see text).
‖ Pulmonary entities under investigation or unclassified.
¶ Diagnosis of chronic GVHD requires biopsy.

APPENDIX D

DERIVATION OF A SEQUENTIAL TEST STATISTIC FOR CENSORED EXPONENTIAL DATA

APPENDIX D

DERIVATION OF A SEQUENTIAL TEST STATISTIC FOR CENSORED EXPONENTIAL DATA

Background – The Sequential Probability Ratio Test

Let $f(.,\theta)$ be the density function for random variable X. According to Neyman and Pearson, the most powerful test of $H_0:\theta=\theta_o$ versus $H_1:\theta=\theta_1$ decides in favor of H_1 or H_0 if $L_n > c_\alpha$ or $L_n < c_\alpha$, respectively, where $L_n = \prod_i^n f(x_i;\theta_1)/f(x_i;\theta_0)$ is the likelihood ratio, and c_α is determined to have the size α. When the sample size is not fixed in advance, further improvement is possible by using Wald's Sequential Probability Ratio Test (SPRT). The SPRT continues to sample as long as $B < L_n < A$ for some constant $B<1<A$, stops sampling and decides in favor of H_1 as soon as $L_n > A$, and stops sampling and decides in favor of H_0 as soon as $L_n < B$.

The usual measures of performance of such a procedure are the error probabilities α and β of rejecting H_0 when $\theta=\theta_0$, and of accepting H_0 when $\theta=\theta_1$, respectively, and the expected sample size $E(N\mid\theta_j)\equiv E_j(N)$. Wald and Wolfowitz showed that among all tests, sequential or not, for which $\Pr_0(\text{reject } H_0)\le\alpha$ and $\Pr_1(\text{reject } H_0)\le\beta$, and for which $E_j(N)$ are finite, j=0,1, the SPRT with error probabilities α and β minimizes $E_0(N)$ and $E_1(N)$. If, in addition, the $x1, x2, ...$ are independent and identically distributed (i.i.d.) with density function $f(x,\theta)$, with monotone likelihood ratio in $\tau(x)$, then any SPRT for testing θ_0 against $\theta_1(>\theta_0)$ has non-decreasing power function.

For the SPRT with error probabilities α and β, the SPRT boundaries are given approximately by $A=(1-\beta)/\alpha$ and $B=\beta/(1-\alpha)$. The operating characteristics of the SPRT are given by $O(\theta,\alpha,\beta,\theta_0,\theta_1)=(A^{h(\theta)}-1)/(A^{h(\theta)}-B^{h(\theta)})$ where $h(\theta)$ is the non-trivial solution to the equation $\int(f(x;\theta_1)/f(x,\theta_2))^{h(\theta)}f(x;\theta)dx=1$.

The formula $E(N;\theta)=[[(1-O(\theta)]\log A+O(\theta)\log B]/E(z;\theta)$ provides the average sample number for an arbitrary θ. The sample size distribution is very highly skewed, $Var(N)\approx[E(N)]^2$. Thus we will consider a truncated test with maximum sample size of N_0 and simulate to obtain the operating characteristics of the test.

Derivation of the SPRT for Uncensored Exponential Survival Times

For example, we wish to construct a sequential test for the composite null hypothesis that the rate of treatment-related mortality (TRM) at 100 days is less than or equal to 30% versus the alternative hypothesis that it is greater than or equal to 50%. For the derivation of the uncensored SPRT, we will require that the type I error of the test be less than 5%, and that the test provide 80% power to reject the null hypothesis under a specified alternative that the true rate is 50%. A maximum sample size of 50 patients will be permitted.

Let us assume that the survival times, $T_1, T_2, ..., T_n$, are completely observed (uncensored) and are i.i.d. with exponential density function $f(T,\theta) = \theta e^{-\theta T}$. These assumptions will be relaxed to incompletely observed data subsequently. In the exponential parameterization, a 100-day survival rate of 70% translates into a mean survival of 0.768 years ($\theta_0 = 1.303$), and 50% translates into a mean survival of 0.395 years ($\theta_1 = 2.532$).

The SPRT is derived with reference to a simple null and alternative hypothesis, in this case, $H_0 : \theta = \theta_o = 1.303$ versus $H_1 : \theta = \theta_1 = 2.532$. However, since the log-likelihood ratio for the exponential, $\log \prod_i^n f(x_i;\theta_1) - \log \prod_i^n f(x_i,\theta_0) = n(\log(\theta_1) - \log(\theta_0)) - (\theta_1 - \theta_0)\sum_i^n T_i$, is a monotone function of $\sum_i^n T_i$, the power of the test is non-decreasing in θ. Thus the SPRT is a one-sided level .05 test of a composite null ($H_0 : \theta \le \theta_o = 1.303$) versus a composite alternative ($H_1 : \theta \ge \theta_o = 1.303$), with power of $1 - \beta = .80$ at the selected alternative $\theta = \theta_1 = 2.532$.

The SPRT can be represented graphically. The continuation region is bounded by two parallel lines with common slope $(\log\theta_0 - \log\theta_1)/(\theta_0 - \theta_1) = 0.541$, and intercepts $\log A/(\theta_0 - \theta_1) = -2.256$ and $\log B/(\theta_0 - \theta_1) = 1.270$, for the lower and upper bounds, respectively. As each individual unit is put on trial and observed to fail, the cumulative sum of failure times, $\sum_i^n T_i$, is recomputed, and plotted against the current sample size, n. When this graph crosses the lower boundary, the null hypothesis is rejected.

The maximum sample size of 50 patients requires that the SPRT be truncated. We choose to truncate the SPRT by declaring that if the test has failed to terminate after 50 patients, that the null hypothesis will be accepted. Since the probability that the untruncated SPRT would reject the null at a sample size of 50 is negligible, it makes little difference how the final boundary value is selected, and this rule is chosen for simplicity.

Derivation of a Modified SPRT for Censored Exponential Data

The assumption of uncensored exponential survival times is flawed. However, we consider it reasonable to assume the hazard for TRM is constant over the first 100 days post-transplant, and

we will restrict our attention to this time interval. Furthermore, it is not practical to conduct a clinical study by putting each individual on trial, and waiting until that individual is observed to fail. We relax our assumptions as follows. Firstly, each individual's time on study will be computed as time from transplant to failure, or to the 100 day time point, whichever comes first. Secondly, we will put individuals on trial as soon as they become available, without waiting for the previous individual to fail.

Let us consider the impact of relaxing these assumptions one at a time. In a fixed sample size trial with uncensored exponential failure times, mean survival time is estimated by the sample mean of the failure times, or total time on study divided by the number of individuals enrolled. When censoring is introduced, the estimate becomes the total time on study divided by the number of observed (non-censored) failures. This suggests that in an exponential SPRT test modified to incorporate censoring, we replace the observed failure times, $T_1, T_2, ..., T_n$, with censored failures times, $x_1, x_2, ..., x_n$, and the current sample size, n, with the number of observed failures, d.

Now we relax the second assumption, and put individuals on trial as soon as they become available, without waiting for the previous individual to fail. Assume that three years are required for accrual of 50 patients to the study, and that the final analysis takes place 100 days after the last patient is entered. Putting all of this together, we propose a modified truncated SPRT, where at any interim time point, s, ranging from 0 to 3 years 100 days, the sum of observed time on study, $\sum_i^n X_i(s)$ is plotted against the number of observed failures, $d(s)$. In practice, monitoring will be scheduled monthly after the start of enrollment to the study. A further modification to the SPRT was to only use the lower boundary for stopping since the primary focus of the monitoring is to protect against unacceptable 100-day TRM rates.

Operating Characteristics of the Modified SPRT Test for Censored Exponential Data

Recall that the uncensored SPRT targeted a drop in survival at Day 100 from 70% to 50%, with type I and II errors of 5% and 20%. Since only the lower boundary is used for monitoring, the continuation region of the test was bounded below by a line with a slope of 0.541 and intercept of –2.256. The effect of truncation is to reduce the power of the test. In order to compensate for this, we raise the lower boundary to make it easier to cross. Under the further assumption of uniform accrual over a three year period, and monthly interim analyses over the course of the study, the operating characteristics of the modified SPRT were obtained from a simulation study. These simulation show that an intercept of –1.741, corresponding to setting parameters α and β to 10% and 15%, result in empirical type I and II error rates of about 5% and 20%.

Table E-1 Operating Characteristics of Sequential Testing Procedures from a Simulation Study with 100,000 Replications

Treatment-Related Mortality (TRM)

True 100-Day Rate	30%	35%	40%	45%	50%
Probability Reject Null	0.07	0.20	0.41	0.66	0.86
Mean Month Stopped	34.5	32.3	28.5	23.5	18.5
Mean # Endpoints in 100 Days	13.8	15.0	15.1	14.0	12.1
Mean # Patients Enrolled	48	45	40	33	26

While the motivation for this testing procedure is largely heuristic rather than theoretical, the simulation results validate the approach. When the true rate of TRM on or before Day 100 was 30%, the test crossed the lower boundary in 7119 of 100,000 replications, for an estimated type I error rate of 7%. When the true rate of TRM on or before Day 100 was 50%, the test failed to cross the boundary in 14226 of 100,000 replications, for an estimated type II error rate of 14%. And on average, the boundary will be crossed at 18.5 months, when 26 patients will be enrolled to the study.

It is interesting to note that the SPRT derived above for exponential failure times with censoring at 100 days, has operating characteristics which are similar to those of a more traditional SPRT, derived for binomial variates with success probability equal to the 100 day failure rate. Using time to failure rather than a simple binary indicator of failure, leads to little improvement in power when failure times are censored relatively soon after entry on study. We speculate that if the constant hazard rate over the first 100 days were high, the exponential test would reject faster than the binomial test, but have not conducted simulation studies to demonstrate this.

APPENDIX E

SUGGESTED PREDNISONE TAPER FOR RESPONDING PATIENTS

APPENDIX E

SUGGESTED PREDNISONE TAPER FOR RESPONDING PATIENTS

Days		Dose
Minimum of 3 days		**2mg/kg/day**
Minimum of 7 days		**1mg/kg/day**
Taper weekly thereafter at the following intervals	**Week 2**	**0.5 mg/kg/day once daily**
	Week 3	**0.25 mg/kg/day once daily**
	Week 4	**0.2 mg/kg/day once daily**
	Week 5	**0.1 mg/kg/day once daily**
	Week 6	**0.1 mg/kg/day every other day**
	Week 7	**Discontinue**

APPENDIX F

REFERENCES

APPENDIX F

REFERENCES

1. Holtan SG, Pasquini M, Weisdorf DJ. Acute graft-versus-host disease: a bench-to-bedside update. Blood 2014;124:363-73.
2. Ziakas PD, Zervou FN, Zacharioudakis IM, Mylonakis E. Graft-versus-host disease prophylaxis after transplantation: a network meta-analysis. PloS one 2014;9:e114735.
3. Weisdorf D, Haake R, Blazar B, et al. Treatment of moderate/severe acute graft-versus-host disease after allogeneic bone marrow transplantation: an analysis of clinical risk features and outcome. Blood 1990;75:1024-30.
4. MacMillan ML, DeFor TE, Weisdorf DJ. The best endpoint for acute GVHD treatment trials. Blood 2010;115:5412-7.
5. MacMillan ML, DeFor TE, Weisdorf DJ. What predicts high risk acute graft-versus-host disease (GVHD) at onset?: identification of those at highest risk by a novel acute GVHD risk score. British journal of haematology 2012;157:732-41.
6. MacMillan ML, Robin M, Harris AC, et al. A Refined Risk Score for Acute Graft-versus-Host Disease that Predicts Response to Initial Therapy, Survival, and Transplant-Related Mortality. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation 2015.
7. Srinivasan R, Daniels J, Fusaro V, et al. Accurate diagnosis of acute graft-versus-host disease using serum proteomic pattern analysis. Experimental hematology 2006;34:796-801.
8. Wang H, Clouthier SG, Galchev V, et al. Intact-protein-based high-resolution three-dimensional quantitative analysis system for proteome profiling of biological fluids. Mol Cell Proteomics 2005;4:618-25.
9. Levine JE, Braun TM, Harris AC, et al. A prognostic score for acute graft-versus-host disease based on biomarkers: a multicentre study. Lancet Haematology 2015;2:e21–e9.
10. Alousi AM, Weisdorf DJ, Logan BR, et al. Etanercept, mycophenolate, denileukin, or pentostatin plus corticosteroids for acute graft-versus-host disease: a randomized phase 2 trial from the Blood and Marrow Transplant Clinical Trials Network. Blood 2009;114:511-7.
11. Bolanos-Meade J, Logan BR, Alousi AM, et al. Phase 3 clinical trial of steroids/mycophenolate mofetil vs steroids/placebo as therapy for acute GVHD: BMT CTN 0802. Blood 2014;124:3221-7; quiz 335.
12. Pidala J, Tomblyn M, Nishihori T, et al. Sirolimus demonstrates activity in the primary therapy of acute graft-versus-host disease without systemic glucocorticoids. Haematologica 2011;96:1351-6.
13. Ho VT, Cutler C, Carter S, et al. Blood and marrow transplant clinical trials network toxicity committee consensus summary: thrombotic microangiopathy after hematopoietic stem cell transplantation. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation 2005;11:571-5.
14. Marty FM, Lowry CM, Cutler CS, et al. Voriconazole and sirolimus coadministration after allogeneic hematopoietic stem cell transplantation. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation 2006;12:552-9.

15. Mielcarek M, Furlong T, Storer BE, et al. Effectiveness and safety of lower dose prednisone for initial treatment of acute graft-versus-host disease: a randomized controlled trial. Haematologica 2015;100:842-8.
16. Cleeland CS, Mendoza TR, Wang XS, et al. Assessing symptom distress in cancer patients: the M.D. Anderson Symptom Inventory. Cancer 2000;89:1634-46.
17. McQuellon RP, Russell GB, Cella DF, et al. Quality of life measurement in bone marrow transplantation: development of the Functional Assessment of Cancer Therapy-Bone Marrow Transplant (FACT-BMT) scale. Bone Marrow Transplant 1997;19:357-68.
18. McHorney CA, Ware JE, Jr., Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31:247-63.
19. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. Med Care 2001;39:800-12.
20. Lawitschka A, Guclu ED, Varni JW, et al. Health-related quality of life in pediatric patients after allogeneic SCT: development of the PedsQL Stem Cell Transplant module and results of a pilot study. Bone Marrow Transplant 2014;49:1093-7.
21. Harris AC, Young R, Devine S, et al. International, Multicenter Standardization of Acute Graft-versus-Host Disease Clinical Data Collection: A Report from the Mount Sinai Acute GVHD International Consortium. Biology of Blood and Marrow Transplantation 2016;22:4 - 10.
22. Jagasia MH, Hildegard TG, Arora M, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group Report. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation 2015;21:389-401

Hartwell MJ, Özbek U, Holler E, et al. An early biomarker algorithm predicts lethal graft-versus-host disease and survival. JCI Insight. 2017 Feb 9;2(3):e89798. doi: 10.1172/jci.insight.89798.

Major-Monfried H, Özbek U, Renteria AS. Biomarkers predict graft-vs-host disease outcomes better than clinical response after one week of treatment. Blood 2016 128:510 (abstract).