

## TRANSPLANTATION

## Failure-free survival after initial systemic treatment of chronic graft-versus-host disease

Yoshihiro Inamoto,<sup>1</sup> Mary E. D. Flowers,<sup>1,2</sup> Brenda M. Sandmaier,<sup>1,2</sup> Sahika Z. Aki,<sup>1</sup> Paul A. Carpenter,<sup>1,3</sup> Stephanie J. Lee,<sup>1,2</sup> Barry E. Storer,<sup>1,4</sup> and Paul J. Martin<sup>1,2</sup>

<sup>1</sup>Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA; and Departments of <sup>2</sup>Medicine, <sup>3</sup>Pediatrics, and <sup>4</sup>Biostatistics, University of Washington School of Medicine, Seattle, WA

## Key Points

- Failure-free survival is a potentially useful, efficient, and robust basis for interpreting results of initial treatment of chronic GVHD.

This study was designed to characterize failure-free survival (FFS) as a novel end point for clinical trials of chronic graft-versus-host disease (GVHD). The study cohort included 400 consecutive patients who received initial systemic treatment of chronic GVHD at our center. FFS was defined by the absence of second-line treatment, nonrelapse mortality, and recurrent malignancy during initial treatment. The FFS rate was 68% at 6 months and 54% at 12 months after initial treatment. Multivariate analysis identified 4 risk factors associated with treatment failure: time interval <12 months from transplantation to initial treatment, patient age ≥60 years, severe involvement of the gastrointestinal tract, liver,

or lungs, and Karnofsky score <80% at initial treatment. Initial steroid doses and the type of initial treatment were not associated with risk of treatment failure. Lower steroid doses after 12 months of initial treatment were associated with long-term success in withdrawing all systemic treatment. FFS offers a potentially useful basis for interpreting results of initial treatment of chronic GVHD. Incorporation of steroid doses at 12 months would increase clinical benefit associated with the end point. Studies using FFS as the primary end point should measure changes in GVHD-related symptoms, activity, damage, and disability as secondary end points. (*Blood*. 2014;124(8):1363-1371)

## Introduction

Chronic graft-versus-host disease (GVHD) occurs in 30% to 50% of patients after allogeneic hematopoietic cell transplantation.<sup>1,2</sup> This complication causes significant late morbidity and mortality and prolongs the duration of immunosuppressive treatment (IST).<sup>3,4</sup> Although novel drugs targeting specific mechanisms have been suggested for treating GVHD,<sup>5</sup> the lack of standardized and validated response criteria remains a major obstacle in designing and interpreting clinical trials of treatment of chronic GVHD.<sup>6</sup> The provisional response criteria were proposed based on expert opinion in the 2005 National Institutes of Health (NIH) Consensus Conference.<sup>6</sup> Results from a recent prospective therapeutic trial showed that response defined according to these criteria correlated with survival in patients with steroid-refractory chronic GVHD,<sup>7</sup> but response did not correlate with survival in a multicenter, observational study of incident and prevalent cases of chronic GVHD.<sup>8</sup>

We have proposed that the absence of systemic treatment change, nonrelapse mortality, and recurrent malignancy could be incorporated into a composite end point to evaluate results of treatment of chronic GVHD.<sup>7</sup> This composite failure-free survival (FFS) end point is more relevant than the individual components, because each of these events should be considered as failure in a clinical trial. The validity of the FFS end point has been evaluated after second-line systemic treatment of chronic GVHD<sup>9</sup> but not after initial systemic treatment.

The current study was designed to (1) determine the FFS rate after initial systemic treatment of patients with moderate or severe chronic GVHD according to NIH criteria, (2) elucidate risk factors associated with treatment failure, and (3) determine whether the dose of prednisone after 6 or 12 months of treatment is associated with long-term success in withdrawing all systemic treatment after resolution of chronic GVHD. Results after initial treatment were compared with those after second-line treatment, and success rates in this retrospective study were compared with those in a previous multicenter prospective study of initial systemic treatment.

## Methods

## Patients and data collection

The study cohort included 400 consecutive relapse-free patients who received initial systemic treatment of moderate or severe chronic GVHD at the Fred Hutchinson Cancer Research Center/Seattle Cancer Care Alliance between 2006 and 2010 after a first allogeneic transplantation. All patients were included regardless of the indication for transplantation, the conditioning regimen, graft source, donor relationship, or HLA matching.

Involved sites and types of treatment at onset of initial systemic treatment of chronic GVHD and treatment change after initial treatment were recorded prospectively via the Long-Term Follow-Up Program through medical

Submitted March 19, 2014; accepted May 27, 2014. Prepublished online as *Blood* First Edition paper, May 29, 2014; DOI 10.1182/blood-2014-03-563544.

The online version of this article contains a data supplement.

There is an Inside *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2014 by The American Society of Hematology

**Table 1. Patient characteristics (N = 400)**

Characteristic	No. (%)
Median patient age at initial treatment, range, years	52 (1-79)
<b>Patient gender</b>	
Male	220 (55)
Female	180 (45)
<b>Donor-patient gender combination</b>	
Female to male	108 (27)
Other	292 (73)
<b>Diagnosis</b>	
Myeloid malignancy	232 (58)
Lymphoid malignancy	154 (39)
Other/nonmalignant	14 (4)
<b>Disease risk at transplantation*</b>	
Low	136 (34)
High	264 (66)
<b>Conditioning regimen</b>	
High dose without TBI	137 (34)
High dose with TBI	81 (20)
Reduced intensity	182 (46)
<b>Graft source</b>	
Bone marrow	46 (12)
Mobilized blood cells	332 (83)
Cord blood	22 (6)
<b>Donor and HLA type</b>	
HLA-matched related	147 (37)
HLA-matched unrelated	155 (39)
HLA antigen or allele-mismatched related	19 (5)
HLA antigen or allele-mismatched unrelated	79 (20)

TBI, total body irradiation.

\*The low-risk category included chronic myeloid leukemia in chronic phase, acute leukemia in first remission, myelodysplastic syndrome without excess blasts, and nonmalignant diseases. The high-risk category included all other diseases and stages.

records from our outpatient clinic and local clinics that provided primary care for patients. Services of the program include (1) evaluation of GVHD patients every 3 to 6 months and close communication with primary providers (~800 visits per year), (2) continuous care of patients requiring close monitoring more than once a month, and (3) telemedicine actively provided to patients and primary physicians by a dedicated team of experts. The values of platelet count, serum total bilirubin, and steroid doses, the NIH global score at onset of initial treatment, and steroid doses at 6 and 12 months after initial treatment were abstracted from medical records. All patients had given written consent allowing the use of medical records for research, and the Institutional Review Board of the Fred Hutchinson Cancer Research Center approved the study. The study was conducted in accordance with the Declaration of Helsinki.

## Definitions

Chronic GVHD was defined according to the NIH consensus criteria.<sup>10</sup> Lung involvement was defined according to the NIH criteria for bronchiolitis obliterans.<sup>11</sup> Liver involvement was defined as an NIH liver score  $\geq 1$ , where serum transaminase, alkaline phosphatase, and bilirubin concentrations were all taken into account. Treatment change was defined as any additional systemic treatment not used for initial treatment of chronic GVHD. Addition of systemic steroid for patients who were not initially treated with steroid was considered as treatment change. An increase in steroid dose in patients who were initially treated with steroid was not considered as treatment change, because temporarily increased steroid doses or resumption of steroid treatment are often necessary during the initial treatment of chronic GVHD.<sup>12</sup>

FFS was defined by the absence of treatment change, nonrelapse mortality, and recurrent malignancy during initial systemic treatment. Recurrent malignancy was defined by hematologic relapse or any unplanned intervention to prevent progression of malignancy in patients with molecular, cytogenetic, flow cytometric, or any other evidence of malignant disease after

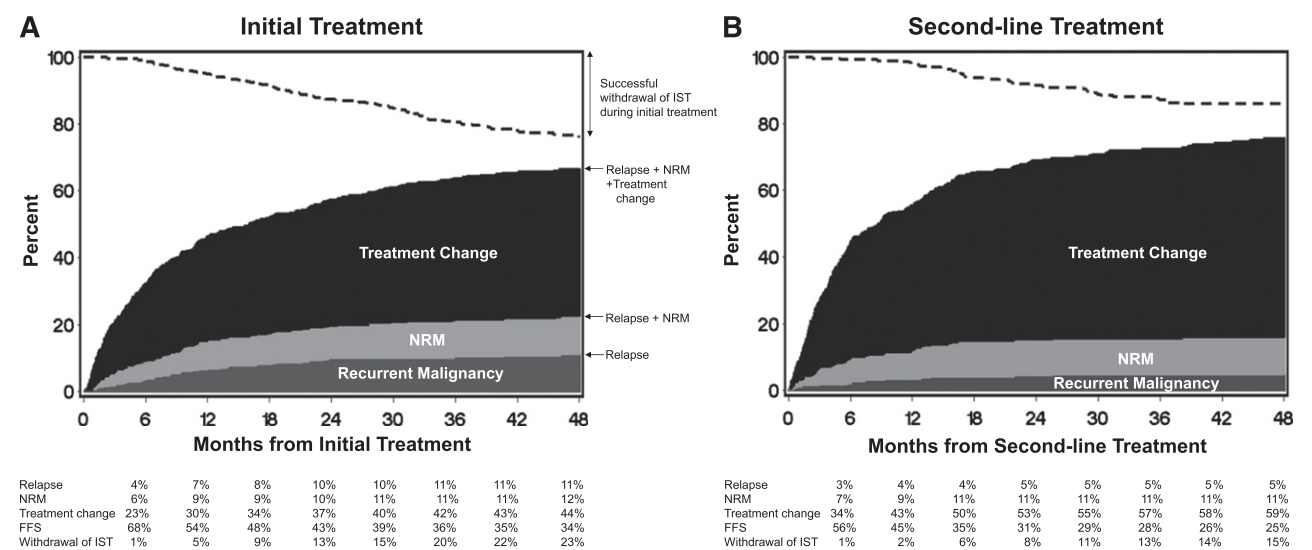
**Table 2. GVHD characteristics at initial systemic treatment (N = 400)**

Characteristic	No. (%)
Median time from transplantation to initial systemic treatment, range, months	6.8 (2.6-58)
<b>Sites involved</b>	
Skin	302 (76)
Eyes	115 (28)
Mouth	363 (91)
Liver	175 (44)
Gastrointestinal tract	
Upper only	78 (20)
Any lower	81 (20)
Lung	14 (4)
Joint or fascia	54 (14)
Genital tract	29 (7)
<b>No. of sites involved</b>	
1 or 2	136 (34)
3	149 (37)
$\geq 4$	115 (29)
<b>NIH global score</b>	
Moderate	228 (57)
Severe	172 (43)
<b>Subcategory of chronic GVHD</b>	
Classic	45 (11)
Overlap	355 (89)
<b>Karnofsky score</b>	
80-100	260 (65)
$< 80$	140 (35)
<b>Platelet count</b>	
$< 100\,000/\mu\text{L}$	105 (26)
$\geq 100\,000/\mu\text{L}$	295 (74)
<b>Serum total bilirubin</b>	
$> 2\text{ mg/dL}$	16 (4)
$\leq 2\text{ mg/dL}$	384 (96)
Progressive onset	39 (10)
Prior grades II-IV acute GVHD	296 (74)
<b>Prednisone-equivalent steroid dose before onset of initial treatment</b>	
None	299 (75)
$< 0.5\text{ mg/kg daily}$	61 (15)
$\geq 0.5\text{ but } < 1.0\text{ mg/kg daily}$	29 (7)
$\geq 1.0\text{ mg/kg daily}$	11 (3)
<b>Initial treatment of chronic GVHD</b>	
Prednisone $\pm$ CNI alone	203 (51)
Prednisone + agents other than CNI	136 (34)
MMF included	96 (24)
Sirolimus included	32 (8)
Other agents included*	21 (5)
Treatment without prednisone	61 (15)
CNI included	55 (14)
MMF included	11 (3)
Sirolimus included	3 ( $< 1$ )
Other agents included†	3 ( $< 1$ )
<b>Prednisone-equivalent steroid dose used for initial treatment</b>	
No prednisone	61 (15)
$< 0.5\text{ mg/kg/daily}$	38 (10)
$\geq 0.5\text{ but } < 1.0\text{ mg/kg/daily}$	105 (26)
$1.0\text{ mg/kg/daily}$	168 (42)
$> 1.0\text{ mg/kg/daily}$	28 (7)
<b>Number of agents used for initial treatment of chronic GVHD</b>	
1	84 (21)
2	216 (54)
$\geq 3$	100 (25)

CNI, calcineurin inhibitor; MMF, mycophenolate mofetil.

\*Other agents: extracorporeal photopheresis (n = 11), methotrexate (n = 5), rituximab (n = 1), thalidomide (n = 1), extracorporeal photopheresis and alemtuzumab (n = 1), infliximab (n = 1), and antithymocyte globulin (n = 1).

†Other agents: extracorporeal photopheresis (n = 1), rituximab (n = 1), and thalidomide (n = 1).



**Figure 1. FFS after systemic treatment of chronic GVHD.** (A) initial treatment, (B) second-line treatment. The dark gray area represents treatment failure due to recurrent malignancy. The light gray area represents treatment failure due to nonrelapse mortality (NRM), and the black area represents treatment failure due to treatment change. The white area represents FFS. The dashed line represents cumulative incidence of successful withdrawal of all systemic IST during initial treatment.

transplantation. An accelerated taper of IST due to evidence of recurrent disease was considered an unplanned intervention and categorized as recurrent malignancy. Permanent withdrawal of all systemic IST was defined as withdrawal of all systemic IST for  $\geq 6$  months without resuming IST.<sup>12,13</sup>

Treatment of chronic GVHD

At our center, chronic GVHD is generally treated by adding prednisone to any other immunosuppressive agents the patient is already taking.<sup>14,15</sup> Prednisone was most often initially administered at a dose of 1 mg/kg once daily for 2 weeks, and the dose was tapered during the subsequent 4 weeks to 1 mg/kg every other day as allowed by improvement in GVHD manifestations. After resolution of reversible manifestations of chronic GVHD, systemic treatment was gradually withdrawn. Decisions to initiate second-line treatment were made at the discretion of the attending physician.

Statistical analysis

Cumulative incidence estimates of recurrent malignancy, nonrelapse mortality, and treatment change as causes of failure during initial treatment were derived, treating each event as a competing risk for the other 2.<sup>16</sup> Permanent withdrawal of all IST during initial treatment after resolution of chronic GVHD was treated as a competing risk for all 3 types of failure. Rates of FFS were estimated by subtracting rates of total failures from 100%. Cox regression models were used to identify risk factors for failure. Factors having a likelihood ratio  $P \leq .05$  for association with failure in univariate testing were included in a multivariate model. A backward elimination procedure was used to exclude risk factors until the  $P$  value of the likelihood ratio test for all remaining risk factors was  $\leq .05$ . Covariates at transplantation included patient age, disease risk, conditioning intensity, graft source, donor relation, HLA matching, and gender matching. Covariates at initial treatment included time from transplantation, involved sites, score of 3 in the gastrointestinal tract, liver, or lungs, number of involved sites, Karnofsky score, NIH global score, classic vs overlap subcategory, thrombocytopenia, hyperbilirubinemia, progressive onset, prior grade II to IV acute GVHD, steroid dose before initial treatment, and type and number of agents used for initial treatment. Prednisone-equivalent steroid doses of 0.3, 0.2, and 0.05 mg/kg daily at 6 months and 0.2, 0.05, and 0 mg/kg daily at 12 months were used as upper limits in evaluating the significance of incorporating prednisone dose in the FFS end point. These doses corresponded to the 75th, 50th, and 25th percentile doses at 6 and 12 months, respectively, among patients initially treated with regimens including steroids. The analysis was carried out as of August 2013.

Results

Characteristics of the study cohort

Patient characteristics are summarized in Table 1. The median age of patients was 52 years (range, 1-79 years). Two hundred eighteen patients (54%) were prepared with high-dose conditioning regimens, and 332 (83%) received mobilized blood cell grafts.

GVHD characteristics at initial systemic treatment are summarized in Table 2. The median time from transplantation to initial systemic treatment of chronic GVHD was 6.8 months (range, 2.6-58 months). The sites most frequently involved at the onset of initial systemic treatment were the mouth (91%) and skin (76%), followed by the liver (44%), gastrointestinal tract (40%), and eyes (28%). One hundred fifteen patients (29%) had involvement of  $\geq 4$  sites, 172 (43%) had a severe NIH global score, 335 (89%) had overlap chronic GVHD, 140 (35%) had a Karnofsky score  $< 80\%$ , 105 (26%) had thrombocytopenia, 16 (4%) had hyperbilirubinemia, and 11 (3%) were receiving prednisone doses  $\geq 1.0$  mg/kg daily immediately before initial treatment. Initial treatment of chronic GVHD included prednisone with or without a calcineurin inhibitor in 203 patients (51%), prednisone with other combinations in 136 (34%), and treatment without prednisone in 61 patients (15%). Initial daily prednisone doses were  $\geq 1.0$  mg/kg in 196 patients (49%),  $\geq 0.5$  but  $< 1.0$  mg/kg in 105 (26%), and  $< 0.5$  mg/kg in 38 (10%). Sixty-one patients (15%) were not initially treated with steroids because they did not tolerate steroids or because GVHD could be adequately managed without steroids. One hundred patients (25%) were treated with regimens including  $\geq 3$  agents. The median follow-up among survivors was 53 months (range, 6-88 months) after initial treatment.

Outcomes after initial and second-line systemic treatment

The FFS rate after initial treatment was 68% (95% confidence interval [CI], 63-72%) at 6 months and 54% (95% CI, 48-58%) at 12 months (Figure 1A). Treatment change was the predominant category of treatment failure. Rates of nonrelapse mortality and recurrent malignancy as causes of treatment failure during initial

**Table 3. Risk factors associated with treatment failure**

Factor	N	Univariate		Multivariate	
		Hazard ratio (95% CI)	P	Hazard ratio (95% CI)	P
Patient age at initial treatment					
0-39	97	1.00 (reference)		1.00 (reference)	
40-49	77	1.05 (0.71-1.55)	.81	1.00 (reference)	
50-59	125	1.07 (0.77-1.49)	.70	1.00 (reference)	
≥60	101	1.43 (1.01-2.01)	.04	1.48 (1.13-1.95)	.004
Female donor to male recipient	108	1.03 (0.79-1.35)	.81		
High-risk disease at transplantation	264	1.00 (0.78-1.29)	.99		
Conditioning regimen					
High dose without TBI	137	1.00 (reference)			
High dose with TBI	81	1.38 (0.97-1.95)	.07		
Reduced intensity	182	1.46 (1.10-1.94)	.009		
Graft source					
Bone marrow	46	1.00 (reference)			
Mobilized blood cells	332	1.22 (0.79-1.87)	.37		
Cord blood	22	0.90 (0.44-1.85)	.78		
Donor and HLA type					
HLA-matched related	147	1.00 (reference)			
HLA-matched unrelated	155	1.12 (0.85-1.47)	.44		
HLA-mismatched	98	0.97 (0.70-1.33)	.84		
Time from transplantation to initial systemic treatment					
<6 mo	166	1.00 (reference)		1.46 (1.05-2.03)	.02
6 to <12 mo	157	0.99 (0.76-1.29)	.92	1.46 (1.05-2.03)	
12 to <24 mo	66	0.74 (0.52-1.07)	.11	1.00 (reference)	
≥24 mo	11	0.31 (0.10-0.99)	.05	1.00 (reference)	
Presence of involvement at initial treatment					
Skin	302	1.13 (0.84-1.51)	.43		
Eyes	115	1.18 (0.91-1.54)	.21		
Mouth	363	1.02 (0.66-1.56)	.94		
Liver	175	0.80 (0.63-1.03)	.08		
Gastrointestinal tract					
No involvement	241	1.00 (reference)			
Upper only	78	1.19 (0.87-1.63)	.27		
Any lower	81	1.35 (1.00-1.83)	.05		
Lungs	14	1.34 (0.71-2.53)	.36		
Joint or fascia	54	0.75 (0.51-1.10)	.14		
Genital tract	29	0.80 (0.49-1.31)	.38		
Score 3 in the gastrointestinal tract, liver, or lungs	54	1.43 (1.02-2.01)	.04	1.43 (1.02-2.02)	.04
No. of sites involved at initial treatment					
1 or 2	136	1.00 (reference)			
3	149	0.80 (0.59-1.07)	.13		
≥4	115	1.10 (0.82-1.49)	.53		
NIH global score at initial treatment					
Moderate	228	1.00 (reference)			
Severe	172	1.09 (0.86-1.40)	.47		
Subcategory of chronic GVHD					
Classic	45	1.00 (reference)			
Overlap	355	1.11 (0.74-1.67)	.61		
Karnofsky score <80%	140	1.32 (1.03-1.70)	.03	1.29 (1.01-1.67)	.04
Thrombocytopenia at initial treatment	105	1.16 (0.89-1.53)	.28		
Hyperbilirubinemia at initial treatment	16	1.13 (0.62-2.06)	.70		
Progressive onset	39	0.79 (0.52-1.21)	.28		
Prior grades II-IV acute GVHD	296	1.10 (0.82-1.46)	.53		
Prednisone dose before onset of initial treatment					
None	299	1.00 (reference)			
<0.5 mg/kg daily	61	1.28 (0.92-1.77)	.14		
≥0.5 mg/kg daily	40	0.92 (0.61-1.38)	.67		
Initial treatment of chronic GVHD					
Prednisone ± calcineurin inhibitor	203	1.00 (reference)			

Treatment failure was defined by the onset of second-line systemic treatment, nonrelapse death, or recurrent malignancy during first-line treatment. TBI, total body irradiation.

\*Other treatments include extracorporeal photopheresis, rituximab, thalidomide, alemtuzumab, infliximab, antithymocyte globulin, and methotrexate. Risk of treatment failure did not differ statistically according to any individual agent.

Table 3. (continued)

Factor	N	Univariate		Multivariate	
		Hazard ratio (95% CI)	P	Hazard ratio (95% CI)	P
Prednisone + others*	136	0.82 (0.62-1.07)	.15		
Treatment without prednisone	61	0.99 (0.69-1.40)	.94		
Prednisone dose used for initial treatment					
No prednisone	61	1.00 (reference)			
<0.5 mg/kg/daily	38	1.05 (0.64-1.72)	.85		
≥0.5 but <1.0 mg/kg/daily	105	0.87 (0.59-1.29)	.49		
≥1.0 mg/kg/daily	196	0.95 (0.67-1.35)	.78		
Number of agents used for initial treatment					
1	84	1.18 (0.86-1.60)	.31		
2	216	1.00 (reference)			
≥3	100	0.88 (0.66-1.18)	.40		

Treatment failure was defined by the onset of second-line systemic treatment, nonrelapse death, or recurrent malignancy during first-line treatment. TBI, total body irradiation.  
\*Other treatments include extracorporeal photopheresis, rituximab, thalidomide, alemtuzumab, infliximab, antithymocyte globulin, and methotrexate. Risk of treatment failure did not differ statistically according to any individual agent.

treatment were similar to each other and much lower than the rate of treatment change. The cumulative incidence of successful withdrawal of all IST after resolution of chronic GVHD during initial treatment was 13% (95% CI, 9-16%) at 24 months and reached 23% (95% CI, 19-28%) at 48 months. Treatment change occurred less frequently and later during initial treatment than during second-line treatment, whereas successful withdrawal of all IST occurred more frequently and earlier during initial treatment than during second-line treatment (Figure 1A-B). Across the observed 4-year time spans, rates of recurrent malignancy were slightly higher during initial treatment than during second-line treatment. Rates of nonrelapse mortality were similar between initial and second-line treatment.

Risk factors associated with treatment failure after initial systemic treatment

In univariate analysis (Table 3), factors associated with risk of treatment failure included patient age, conditioning regimen, interval time from transplantation to initial systemic treatment, lower gastrointestinal involvement at initial treatment, score 3 in the gastrointestinal tract, liver, or lungs at initial treatment, and

Karnofsky score at initial treatment. Initial steroid doses, type of initial treatment, and the number of agents used for initial treatment were not statistically associated with risk of treatment failure. No individual agent was associated with a statistically significant difference in the risk of treatment failure. In multivariate models, 4 risk factors remained statistically significant: onset of initial treatment of chronic GVHD within 12 months after transplantation, patient age ≥60 years, score 3 in the gastrointestinal tract, liver, or lungs, and Karnofsky score <80% at initial treatment. All 4 risk factors were associated with increased risk of nonrelapse mortality. Karnofsky score <80% and score 3 in the gastrointestinal tract, liver, or lungs at initial treatment were associated with increased risk of recurrent malignancy. None of the 4 risk factors was associated with treatment change.

Treatment failure rates according to risk groups

Forty-one patients (10%) had none of the 4 identified risk factors, 158 (40%) had 1 risk factor, 147 (37%) had 2 risk factors, 50 (13%) had 3 risk factors, and 4 (1%) had 4 risk factors. Rates of treatment failure were similar among groups with 0, 1, or 2 risk factors, and these 346 patients (86%) were categorized as standard risk. Rates of

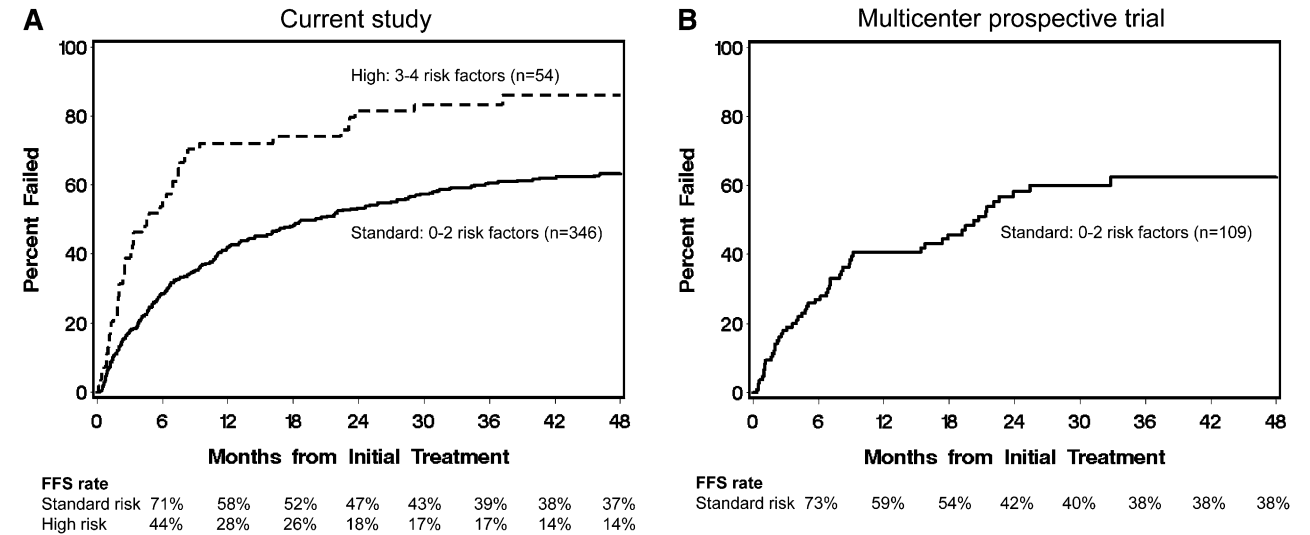
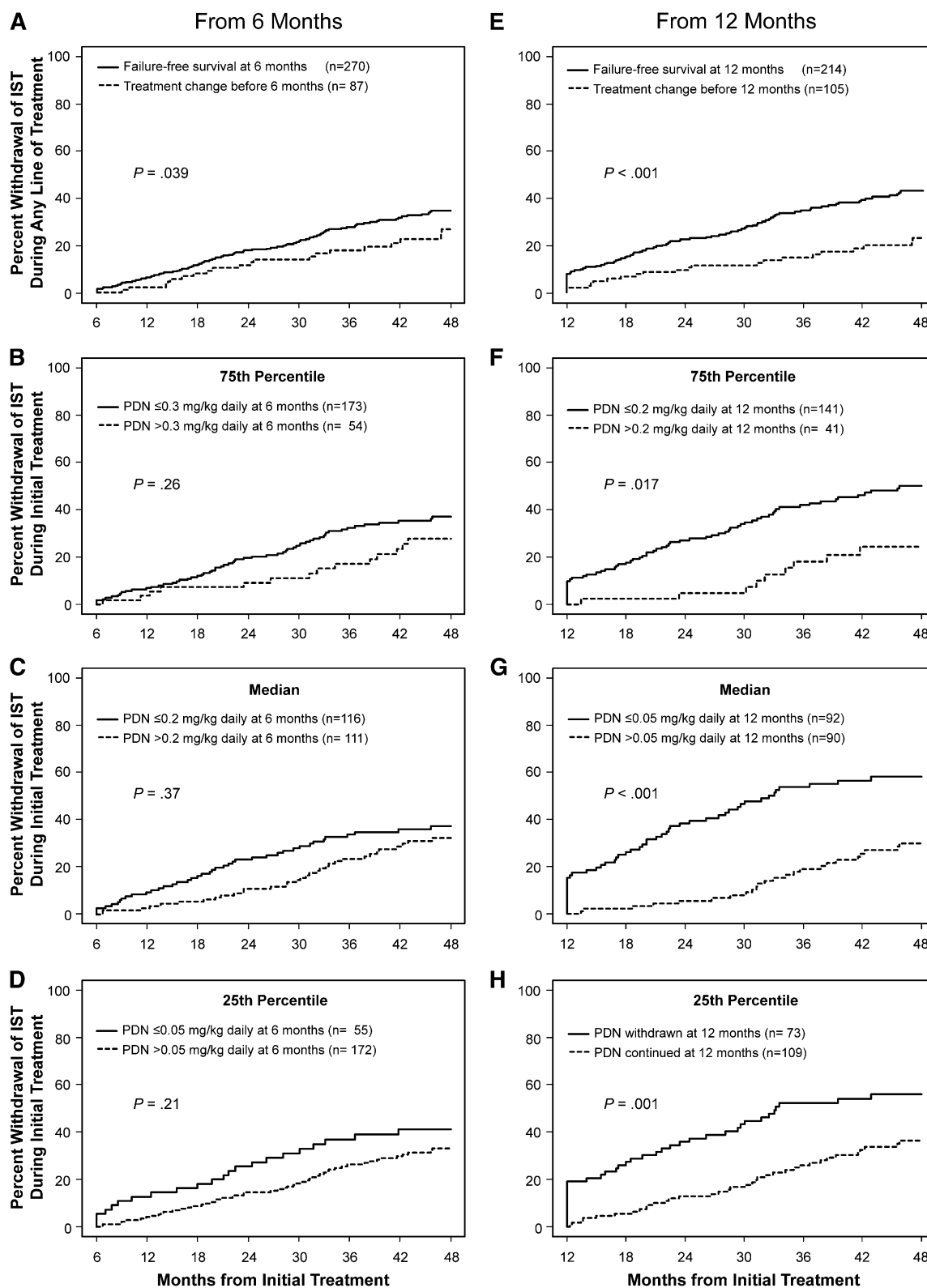


Figure 2. Treatment failure rates according to risk groups. (A) current study, (B) prior multicenter prospective trial. Standard risk, 0 to 2 risk factors; high risk, 3 to 4 risk factors. Risk factors included time interval <12 months from transplantation to initial treatment, patient age ≥60 years, severe involvement (score 3) of the gastrointestinal tract, liver, or lungs, and Karnofsky score <80% at initial treatment. Results for the high-risk group were not plotted in B, because only 2 patients were included in this group. These 2 patients had early treatment failure at 15 and 19 days after initial treatment, respectively.



**Figure 3. Long-term success in withdrawing all immunosuppressive treatment.** Results were analyzed according to the presence or absence of (A,E) prior treatment change and prednisone (PDN) doses at (B-D) 6 and (F-H) 12 months after initial treatment. Withdrawal of all immunosuppressive treatment after resolution of GVHD among patients without recurrent malignancy was counted during any line of treatment in A and E and during initial treatment among patients initially treated with steroids in B-D and F-H. Patients with recurrent malignancy before the landmark were excluded in all analyses, and patients with treatment change before the landmark were excluded in the analyses shown in B-D and F-H.

Table 4. Success rates at 12 months according to risk groups at initial treatment

Risk group*	No.	Percentage	FFS	FFS + PDN ≤ 0.2 mg/kg/day	FFS + PDN ≤ 0.05 mg/kg/day	FFS + no PDN
All patients	400	100	0.54	0.42	0.27	0.22
Standard	346	86	0.58	0.45	0.29	0.23
High	54	14	0.28	0.22	0.15	0.11

PDN, prednisone.  
\*Standard risk, 0 to 2 risk factors; high risk, 3 to 4 risk factors. Risk factors included time interval <12 months from transplantation to initial treatment, patient age ≥60 years, severe involvement (score 3) of the gastrointestinal tract, liver, or lungs, and Karnofsky score <80% at initial treatment.

treatment failure were higher among 54 patients (14%) with 3 or 4 risk factors compared with standard-risk patients, and they were categorized as high risk. FFS rates in the standard and high-risk groups were 71% and 44%, respectively, at 6 months and 58% and 28%, respectively, at 12 months (Figure 2A).

Long-term success in withdrawing all immunosuppressive treatment and survival according to failure-free status and steroid doses at 6 and 12 months after initial treatment

Among 357 and 319 patients without recurrent malignancy at 6 and 12 months after initial systemic treatment, respectively, treatment change before 6 or 12 months was associated with lower subsequent rates of success in withdrawing all IST during any line of treatment ( $P = .039$  and  $< .001$ , respectively) (Figure 3A,E). Landmark analyses at 6 and 12 months did not show statistically significant differences in overall survival according to the presence or absence of prior treatment change (supplemental Figure 1A,E, available on the Blood Web site).

Among 227 patients initially treated with steroid and having FFS at 6 months, the median daily prednisone dose was 0.19 mg/kg (range, 0-1.6 mg/kg) at 6 months. Among 182 patients initially treated with steroid and having FFS at 12 months, the median daily prednisone dose was 0.049 mg/kg (range, 0-1.3 mg/kg) at 12 months. Steroid dose information was not available in 1 patient at 6 and 12 months, and this patient was excluded from the analysis. The cumulative incidence of IST withdrawal after resolution of chronic GVHD during initial treatment did not show statistically significant differences according to steroid thresholds at 6 months (Figure 3B-D). In contrast, the cumulative incidence of IST withdrawal was higher among patients taking prednisone-equivalent steroid doses  $\leq 0.2$  ( $n = 141$ ; 77.5%),  $\leq 0.05$  ( $n = 92$ ; 50.6%), or 0 mg/kg daily ( $n = 73$ ; 40.1%) at 12 months compared with those taking, respectively, higher doses ( $P = .017$ ,  $< .001$  and  $< .001$ , respectively) (Figure 3F-H). The cumulative incidence of IST withdrawal at 48 months after initial treatment was 50% for those with doses  $\leq 0.2$  mg/kg daily, 58% for those with doses  $\leq 0.05$  mg/kg daily, and 56% for those who were taking no prednisone at 12 months. Subsequent overall survival was slightly better among patients taking doses  $\leq 0.3$  or  $\leq 0.2$  mg/kg/day at 6 months compared with those taking respectively higher doses (supplemental Figure 1B-C). Differences were not statistically significant according to steroid doses at 12 months (supplemental Figure 1F-H).

Success rates at 12 months according to risk groups

Table 4 summarizes success rates with and without steroid dose limits at 12 months according to risk groups. The FFS rate was 0.54 at 12 months among all patients. As stated above, proportions of failure-free patients who met the dose limits of prednisone  $\leq 0.2$ ,  $\leq 0.05$ , and 0 mg/kg daily were 77.5%, 50.6%, and 40.1% at 12 months, respectively. When these dose limits were each incorporated as an additional criterion by multiplication, success rates

decreased to 0.42, 0.27, and 0.22, respectively. Patterns were similar in each of the individual risk groups.

Comparison of FFS rates between the current study and a previously performed multicenter prospective trial

FFS rates were calculated using the dataset of a previous multicenter prospective trial of initial systemic treatment.<sup>12</sup> Forty patients included in the current study were excluded from the prospective dataset so that no patients overlapped between the cohorts. Among 111 patients in the multicenter prospective study, 109 were categorized as standard risk. FFS rates were closely similar to those for the standard-risk patients in the current study (Figure 2B). Only 2 patients in the prospective study were categorized as high risk, in part because hospitalized patients and those with bronchiolitis obliterans were excluded. These 2 patients had early treatment failure at 15 and 19 days after initial treatment, respectively. Success rates incorporating prednisone doses were not compared, because information regarding prednisone doses at 12 months was available for only a small proportion of patients in the prospective trial.

Discussion

This study provides important benchmarks for designing future clinical trials. First, we defined FFS as a novel composite end point for clinical trials and characterized the FFS outcomes after initial systemic treatment of moderate or severe chronic GVHD. Second, with treatment change added to nonrelapse mortality and recurrent malignancy in the composite end point, the event rate increased substantially, increasing the opportunity to demonstrate improved results in future studies. Third, we identified 4 objective risk factors associated with treatment failure and defined a group of patients with high risk of treatment failure. Fourth, we found that lower steroid doses among patients with FFS at 12 months were associated with successful withdrawal of all systemic treatment after resolution of chronic GVHD, an unequivocal long-term clinical benefit. Finally, we report success rates accounting both for risk groups and steroid dose limits at 12 months to make the end point more meaningful and to enhance the clinical benefit associated with the end point.

This study included consecutive contemporary patients with moderate or severe chronic GVHD at our center to approximate the characteristics of patients who could be enrolled in future clinical trials of initial treatment of chronic GVHD. Notably, treatment failure was not associated with initial prednisone dose or type of initial treatment, and an independent cohort from a multicenter prospective study showed similar success rates. Thus, the results support the use of FFS as a robust benchmark that could be used for designing and interpreting future phase 2 trials of initial treatment of chronic GVHD.

Early onset of chronic GVHD, low Karnofsky performance, older patient age, and a score of 3 in the gastrointestinal tract, liver, or lungs

were the risk factors associated with increased risk of treatment failure in this study. These risk factors have been associated with mortality after initial treatment.<sup>3,15,17,18</sup> Because these risk factors are objective and easy to evaluate at any transplant center, they are suitable for use in multicenter studies. In a multicenter observational study including both incident and prevalent cases,<sup>19</sup> the risk of mortality was higher in patients with overlap subtype than in those with classic subtype. The proportion of patients with overlap manifestations was 89% among incident cases in the multicenter study,<sup>19</sup> identical to the proportion from our study. In comparing the results between initial and second-line treatment, major differences in outcomes included less frequent treatment change and more frequent and earlier withdrawal of all IST after initial treatment than after second-line treatment, suggesting that chronic GVHD is more responsive during initial treatment as compared with second-line treatment.

In the absence of standardized and validated criteria for measuring treatment response, steroid doses at the time of end point assessment might increase the clinical benefit associated with the end point. Unlike second-line treatment, incorporation of an upper limit of steroid dose in the 6-month FFS end point did not show statistically significant association with successful withdrawal of IST.<sup>9</sup> This difference could be explained by the possibility that evaluation after 6 months of initial treatment is too short to predict long-term outcomes. Some manifestations of chronic GVHD such as sclerosis improve very slowly, and assessment of these manifestations at time points later than 6 months after treatment has been recommended.<sup>20,21</sup> Response at 6 months did not correlate with subsequent development of tolerance,<sup>22</sup> and conflicting results have been reported for the association of response at 6 months with subsequent mortality.<sup>7,8</sup> Although assessment at 6 months is “quick and convenient,” our results support the use of an assessment that incorporates prednisone doses after 12 months of initial treatment. Success with this end point in a phase 2 study would increase the probability that subsequent phase 3 trials will demonstrate differences in successful withdrawal of all systemic treatment as a long-term goal demonstrating unequivocal clinical benefit.

The use of FFS as a primary end point in chronic GVHD treatment trials has several caveats. First, the FFS end point does not give any direct information about changes in GVHD-related symptoms, activity, damage, or disability. Studies using FFS as the primary end point should evaluate all of these measures as secondary end points. Second, results with this end point require careful interpretation in nonblinded trials. The composite end point includes objective components, including death, recurrent malignancy, and the fact of treatment change, but decisions to change treatment are controlled by providers and can be motivated by a variety of factors, including the presence of chronic GVHD manifestations that are progressing, persisting without improvement, or improving more slowly than desired. Changes may also be motivated by toxicity, inconvenience, financial burden, or the availability of a newer alternative that is more attractive to the patient or physician. In nonblinded clinical trials, intentionally delayed treatment changes could increase the risk of false-positive results,

whereas premature treatment changes could increase the risk of false-negative results. Controlled designs with blinding should be used whenever feasible to mitigate these risks when FFS is used as the primary end point. Guidelines for tapering steroid doses and clear definitions of appropriate indications for treatment change could enhance the objectivity of this end point. Lastly, the benchmark rates from this study might not apply to pediatric patients or to patients with nonmalignant diseases, because the study cohort included few such patients.

To date, 6 phase 3 randomized trials have explored whether initial therapy was improved by adding an immunosuppressive agent to conventional corticosteroid treatment,<sup>12,18,23-26</sup> but no study showed superiority of the investigational arm. Therefore, robust and efficient phase 2 studies are needed to identify promising drugs that could be tested in phase 3 studies. FFS serves as a potentially useful basis for designing such trials and for interpreting results of initial treatment of chronic GVHD. Incorporation of a steroid dose limit at 12 months after initial treatment as an additional criterion would increase the degree of clinical benefit associated with FFS. In the future, it should be possible to replace FFS with a standardized and validated measure of response as the primary end point to measure clinical benefit in a more direct way.

## Acknowledgments

We thank Dr Morgani Rodrigues for assistance with some of the data abstraction from the medical records.

This work was supported by NIH grants CA163438, CA18029, CA78902, and CA118953 from the National Cancer Institute. The Chronic GVHD Consortium (grant U54 CA163438) is a part of the NIH Rare Diseases Clinical Research Network, supported through collaboration between the NIH Office of Rare Diseases Research at the National Center for Advancing Translational Science, the National Cancer Institute, and the Fred Hutchinson Cancer Research Center.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

## Authorship

Contribution: Y.I., M.E.D.F., and P.J.M. designed the study, collected and analyzed data, and wrote the paper; B.E.S. performed the statistical analysis and wrote the paper; B.M.S., S.Z.A., P.A.C., and S.J.L. collected data and wrote the paper; and all authors critically revised the manuscript for important intellectual content and approved the manuscript for publication.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Yoshihiro Inamoto, Fred Hutchinson Cancer Research Center, D5-290, 1100 Fairview Ave N, Seattle, WA 98109; e-mail: yinamoto@fhcrc.org.

## References

- Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2003;9(4):215-233.
- Socié G, Schmoor C, Bethge WA, et al; ATG-Fresenius Trial Group. Chronic graft-versus-host disease: long-term results from a randomized trial on graft-versus-host disease prophylaxis with or without anti-T-cell globulin ATG-Fresenius. *Blood*. 2011;117(23):6375-6382.
- Stewart BL, Storer B, Storek J, et al. Duration of immunosuppressive treatment for chronic graft-versus-host disease. *Blood*. 2004;104(12):3501-3506.
- Wingard JR, Majhail NS, Brazauskas R, et al. Long-term survival and late deaths after allogeneic hematopoietic cell transplantation. *J Clin Oncol*. 2011;29(16):2230-2239.
- Blazar BR, Murphy WJ, Abedi M. Advances in graft-versus-host disease biology and therapy. *Nat Rev Immunol*. 2012;12(6):443-458.
- Pavletic SZ, Martin P, Lee SJ, et al; Response Criteria Working Group. Measuring therapeutic response in chronic graft-versus-host disease: National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: IV.



- Response Criteria Working Group report. *Biol Blood Marrow Transplant*. 2006;12(3):252-266.
7. Olivieri A, Cimminiello M, Corradini P, et al. Long-term outcome and prospective validation of NIH response criteria in 39 patients receiving imatinib for steroid-refractory chronic GVHD. *Blood*. 2013;122(25):4111-4118.
  8. Inamoto Y, Martin PJ, Chai X, et al. Chronic GVHD Consortium. Clinical benefit of response in chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2012;18(10):1517-1524.
  9. Inamoto Y, Storer BE, Lee SJ, et al. Failure-free survival after second-line systemic treatment of chronic graft-versus-host disease. *Blood*. 2013;121(12):2340-2346.
  10. Filipovich AH, Weisdorf D, Pavletic S, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. *Biol Blood Marrow Transplant*. 2005;11(12):945-956.
  11. Chien JW, Duncan S, Williams KM, Pavletic SZ. Bronchiolitis obliterans syndrome after allogeneic hematopoietic stem cell transplantation—an increasingly recognized manifestation of chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2010;16(1 Suppl):S106-S114.
  12. Martin PJ, Storer BE, Rowley SD, et al. Evaluation of mycophenolate mofetil for initial treatment of chronic graft-versus-host disease. *Blood*. 2009;113(21):5074-5082.
  13. Inamoto Y, Flowers ME, Lee SJ, et al. Influence of immunosuppressive treatment on risk of recurrent malignancy after allogeneic hematopoietic cell transplantation. *Blood*. 2011;118(2):456-463.
  14. Martin PJ, Gilman AL. Front line treatment of chronic GVHD. In: Vogelsang G, Pavletic S, eds. *Chronic Graft Versus Host Disease: Interdisciplinary Management*. New York: Cambridge University Press; 2009:124-133.
  15. Vigorito AC, Campregher PV, Storer BE, et al; National Institutes of Health. Evaluation of NIH consensus criteria for classification of late acute and chronic GVHD. *Blood*. 2009;114(3):702-708.
  16. Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Stat Med*. 1999;18(6):695-706.
  17. Arora M, Klein JP, Weisdorf DJ, et al. Chronic GVHD risk score: a Center for International Blood and Marrow Transplant Research analysis. *Blood*. 2011;117(24):6714-6720.
  18. Sullivan KM, Witherspoon RP, Storb R, et al. Prednisone and azathioprine compared with prednisone and placebo for treatment of chronic graft-v-host disease: prognostic influence of prolonged thrombocytopenia after allogeneic marrow transplantation. *Blood*. 1988;72(2):546-554.
  19. Pidala J, Vogelsang G, Martin P, et al. Overlap subtype of chronic graft-versus-host disease is associated with an adverse prognosis, functional impairment, and inferior patient-reported outcomes: a Chronic Graft-versus-Host Disease Consortium study. *Haematologica*. 2012;97(3):451-458.
  20. Flowers ME, Apperley JF, van Besien K, et al. A multicenter prospective phase 2 randomized study of extracorporeal photopheresis for treatment of chronic graft-versus-host disease. *Blood*. 2008;112(7):2667-2674.
  21. Greinix HT, van Besien K, Elmaagacli AH, et al; UVADEX Chronic GVHD Study Group. Progressive improvement in cutaneous and extracutaneous chronic graft-versus-host disease after a 24-week course of extracorporeal photopheresis—results of a crossover randomized study. *Biol Blood Marrow Transplant*. 2011;17(12):1775-1782.
  22. Martin PJ, Storer BE, Carpenter PA, et al. Comparison of short-term response and long-term outcomes after initial systemic treatment of chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2011;17(1):124-132.
  23. Koc S, Leisenring W, Flowers ME, et al. Thalidomide for treatment of patients with chronic graft-versus-host disease. *Blood*. 2000;96(12):3995-3996.
  24. Arora M, Wagner JE, Davies SM, et al. Randomized clinical trial of thalidomide, cyclosporine, and prednisone versus cyclosporine and prednisone as initial therapy for chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2001;7(5):265-273.
  25. Koc S, Leisenring W, Flowers ME, et al. Therapy for chronic graft-versus-host disease: a randomized trial comparing cyclosporine plus prednisone versus prednisone alone. *Blood*. 2002;100(1):48-51.
  26. Gilman AL, Schultz KR, Goldman FD, et al. Randomized trial of hydroxychloroquine for newly diagnosed chronic graft-versus-host disease in children: a Children's Oncology Group study. *Biol Blood Marrow Transplant*. 2012;18(1):84-91.