

## TRANSPLANTATION

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

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

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

## Key Points

- Relapse-free survival without treatment change can form the basis of the primary end point in studies of chronic graft-versus-host disease.
- Steroid doses at the time of assessment should be taken into account in treatment studies of chronic graft-versus-host disease.

This study attempted to characterize causes of treatment failure, identify associated prognostic factors, and develop shorter-term end points for trials testing investigational products or regimens for second-line systemic treatment of chronic graft-versus-host disease (GVHD). The study cohort (312 patients) received second-line systemic treatment of chronic GVHD. The primary end point was failure-free survival (FFS) defined by the absence of third-line treatment, nonrelapse mortality, and recurrent malignancy during second-line treatment. Treatment change was the major cause of treatment failure. FFS was 56% at 6 months after second-line treatment. Lower steroid doses at 6 months correlated with subsequent withdrawal of immunosuppressive treatment. Multivariate analysis showed that high-risk disease at transplantation, lower gastrointestinal involvement at second-line treatment, and severe NIH global score at second-line treatment were associated with increased risks of treatment failure. These three factors were used to define risk groups, and success rates at 6 months were calculated for each risk group either without or with various steroid dose limits at 6 months as an additional criterion of success. These success rates

could be used as the basis for a clinically relevant and efficient shorter-term end point in clinical studies that evaluate agents for second-line systemic treatment of chronic GVHD. (*Blood*. 2013;121(12):2340-2346)

## Introduction

Nearly 50% of patients with chronic graft-versus-host disease (GVHD) receive second-line systemic treatment within the first year after initial systemic treatment because of inadequate control of GVHD,<sup>1,2</sup> but no standard second-line treatment has been established. Many different agents have been used, and despite high response rates reported in the literature, actual experience with available agents tends to be unsatisfying.<sup>3,4</sup> New therapies for chronic GVHD are likely to be tested initially for second-line treatment.<sup>5-7</sup> The lack of validated and standardized response criteria represents a major obstacle in designing clinical trials and interpreting the results.<sup>8-11</sup> In a review of 60 published studies of second-line treatment of chronic GVHD during the past 20 years,<sup>6</sup> only one report clearly described a formal statistical hypothesis based on a historical success rate.<sup>12</sup> The absence of data that could be used as a reliable and meaningful basis for the null hypothesis has made it impossible to design robust clinical trials with prespecified statistical considerations.<sup>6</sup>

In clinical trials testing investigational products for treatment of chronic GVHD, causes of failure after second-line treatment can be categorized into 3 types of events: nonrelapse mortality, recurrent malignancy, and systemic treatment change (ie, initiation of third-

line systemic treatment).<sup>5</sup> We hypothesized that failure-free survival (FFS), defined as the absence of these failures, is a meaningful clinical metric that could be used as a shorter-term success end point for clinical trials. Furthermore, we hypothesized that incorporation of an upper limit of the steroid dose at the time of assessment could be used as an additional criterion to enhance the clinical benefit associated with FFS by indicating that GVHD has been well controlled and by decreasing the risk of steroid-related adverse effects.<sup>6,9</sup> We also hypothesized that lower steroid doses at the time of end point assessment may be associated with higher rates of subsequent successful withdrawal of all immunosuppressive treatment after resolution of GVHD.

This study had 4 aims: (1) to characterize causes of treatment failure after second-line treatment of chronic GVHD in consecutive patients representing those who would likely be eligible for future phase II trials, (2) to elucidate prognostic factors associated with treatment failure, (3) to compare rates of withdrawal of all immunosuppressive treatment during second-line treatment according to steroid doses at 6 months, and (4) to provide shorter-term success rates either without or with an upper limit of the steroid dose at 6 months. We propose that these success rates could be used to design and evaluate future phase II studies of second-line treatment of chronic GVHD.

Submitted November 2, 2012; accepted January 2, 2013. Prepublished online as *Blood* First Edition paper, January 15, 2013; DOI 10.1182/blood-2012-11-465583.

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.

© 2013 by The American Society of Hematology

## Patients and methods

### Patients and data collection

Between January 2001 and February 2011, 425 consecutive relapse-free patients at the Fred Hutchinson Cancer Research Center/Seattle Cancer Care Alliance received second-line systemic treatment of chronic GVHD because of either progression or lack of improvement. To mimic the characteristics of patients likely to be enrolled in future phase II trials of second-line treatment of chronic GVHD, 312 of these 425 patients who met all of the following criteria were selected for inclusion in the study cohort: (1) patients who had already received systemic steroid treatment of chronic GVHD at a prednisone-equivalent dose of at least 0.5 mg/kg per day before second-line treatment, (2) patients who were under systemic immunosuppressive treatment when second-line treatment was started, and (3) patients who received second-line treatment because of progressive GVHD manifestations after at least 1 week of initial treatment or because of lack of improvement after at least 2 weeks of initial treatment. Patients were enrolled regardless of the indication for transplantation, the conditioning regimen, graft source, donor relationship, or HLA-matching between the donor and recipient. The following patients were excluded: (1) patients who did not receive prednisone-equivalent steroid dose  $\geq 0.5$  mg/kg per day before second-line treatment ( $n = 64$ ), (2) patients who were not taking any systemic treatment at the start of second-line treatment ( $n = 35$ ), (3) patients with progressive GVHD who started second-line treatment after less than 7 days of initial treatment ( $n = 2$ ) or after a very rapid taper of initial prednisone treatment ( $n = 2$ ), and (4) patients with lack of improvement who started second-line treatment after less than 14 days of initial treatment ( $n = 10$ ).

Involved sites, types of treatment, dates, and the reasons for change of treatment (progression or lack of improvement) after the initial systemic treatment of chronic GVHD were recorded prospectively.<sup>1</sup> The platelet count, serum total bilirubin, steroid doses, and the National Institutes of Health (NIH) global score of chronic GVHD immediately before second-line treatment were collected from medical records. Steroid doses at 6 months were also recorded as part of the assessment after second-line treatment. All patients gave written consent allowing the use of medical records for research in accordance with the Declaration of Helsinki, and the Institutional Review Board approved the study.

### Definitions

Chronic GVHD was defined by the NIH consensus criteria.<sup>13</sup> Lung involvement was defined according to the NIH criteria for bronchiolitis obliterans.<sup>14</sup> Liver involvement was defined as an NIH liver score  $\geq 1$ , in which serum transaminase, alkaline phosphatase, and bilirubin concentrations were all taken into account. Second-line treatment was defined as any additional systemic treatment not used for initial treatment of chronic GVHD. An increase in steroid dose was not considered as second-line treatment, since temporarily increased steroid doses and resumption of steroid treatment are often necessary during the initial treatment of chronic GVHD.<sup>15</sup> This definition of treatment change corresponds to the anticipated situation of future patients who would be candidates for a clinical trial testing an investigational product for treatment of steroid-resistant or steroid-refractory chronic GVHD. Similarly, third-line treatment was defined as any additional systemic treatment not used for second-line treatment of chronic GVHD; an increase in steroid dose was not considered as third-line treatment.

FFS was defined as the absence of third-line treatment, nonrelapse mortality, and recurrent malignancy during second-line treatment. Recurrent malignancy was defined as hematologic relapse or any unplanned intervention to prevent progression of malignancy in patients with molecular, cytogenetic, flow cytometric, or any other kind of evidence of malignant disease after transplantation. An accelerated taper of immunosuppressive treatment because of evidence of recurrent disease was considered as an unplanned intervention and was categorized as recurrent malignancy. Withdrawal of all systemic immunosuppressive treatment was defined as described previously.<sup>15</sup>

### Treatment of chronic GVHD

At our center, initial systemic treatment of chronic GVHD is generally the addition of prednisone to any other immunosuppressive agents the patient is

already taking.<sup>16,17</sup> Prednisone was most often administered initially 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 or subsequent 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 second-line treatment were derived, treating each event as a competing risk for the other two.<sup>18</sup> Successful withdrawal of all immunosuppressive treatment during second-line treatment after resolution or improvement of all reversible manifestations of 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  $P$  value  $\leq .05$  for association with failure in univariate testing were included in a multivariate Cox regression 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  $\leq .05$ . Success rates incorporating steroid doses at 6 months were calculated by multiplying FFS rates at 6 months and the proportion of patients taking the specified steroid doses at 6 months among those with FFS at 6 months. We selected graded prednisone-equivalent steroid doses of 0.3, 0.2, and 0.1 mg/kg per day as upper limits for these analyses. A dose of 0.3 mg/kg per day at 6 months approximates the 75th percentile dose (0.33 mg/kg per day). A dose of 0.2 mg/kg per day approximates the median at 6 months (0.22 mg/kg per day) and can be well tolerated for long periods of time in many patients, since prednisone doses below 0.5 mg/kg every other day are well tolerated in our experience.<sup>16,19,20</sup> A dose of 0.10 mg/kg per day approximates the 25th percentile dose (0.09 mg/kg per day), and this dose also represents a desirable therapeutic goal to minimize steroid-induced complications in patients with rheumatologic diseases<sup>21</sup> or chronic GVHD.<sup>22,23</sup> The analysis was carried out in August 2012.

## Results

### Patient and GVHD characteristics

Patient characteristics are summarized in Table 1. The median age of patients was 48 years (range, 1-73 years). Two hundred ten patients (67%) were prepared with high-intensity conditioning regimens, and 265 (85%) received mobilized blood cell grafts. The median time from transplantation to initial systemic treatment of chronic GVHD was 5.4 months (range, 2.2-29 months).

GVHD characteristics are summarized in Table 2. The median time from initial systemic treatment to second-line treatment was 6.3 months (range, 0.2-61 months). The sites most frequently involved at the onset of second-line treatment were the skin (72%) and mouth (71%). The NIH organ severity scores in each site are shown in supplementary Figure S1. Ninety-one patients (29%) had involvement of 4 or more sites, 128 (42%) had severe NIH global score, 54 (18%) had thrombocytopenia, 23 (7%) had hyperbilirubinemia, and 52 (17%) were receiving prednisone doses  $\geq 1.0$  mg/kg per day immediately before second-line treatment. The most frequently used second-line systemic treatments were mycophenolate mofetil ( $n = 95$ ), tacrolimus ( $n = 63$ ), and sirolimus ( $n = 57$ ). Thirty patients had second-line treatment within 1 month after initial treatment because of unequivocal progression in skin ( $n = 7$ ), liver ( $n = 6$ ), gastrointestinal tract + diarrhea ( $n = 4$ ), skin + eyes + mouth ( $n = 1$ ), skin + liver + diarrhea ( $n = 1$ ), or liver + diarrhea ( $n = 1$ ) or lack of improvement in skin ( $n = 2$ ), liver ( $n = 4$ ), skin + liver ( $n = 1$ ), or diarrhea ( $n = 3$ ). All 30 patients had moderate or severe global score and had received

**Table 1. Patient characteristics (N = 312)**

Characteristic	No. (%)
Median patient age at second-line treatment, range, y	48 (1-73)
Patient gender	
Male	175 (56)
Female	137 (44)
Donor-patient gender combination	
Female to male	89 (29)
Other	223 (71)
Diagnosis	
Myeloid malignancy	208 (67)
Lymphoid malignancy	85 (27)
Other/nonmalignant	19 (6)
Disease risk at transplantation*	
Low	131 (42)
High	181 (58)
Conditioning regimen	
High intensity without TBI	135 (43)
High intensity with TBI	75 (24)
Reduced intensity	102 (33)
Graft source	
Bone marrow	35 (11)
Mobilized blood cells	265 (85)
Umbilical cord blood	12 (4)
Donor and HLA type	
HLA-matched related	130 (42)
HLA-matched unrelated	118 (38)
HLA antigen or allele-mismatched related	9 (3)
HLA antigen or allele-mismatched unrelated	55 (18)
Median time from transplantation to initial systemic treatment, months (range)	5.4 (2.2-29)
Initial systemic treatment	
Prednisone + CNI	201 (64)
Prednisone + CNI + mycophenolate mofetil	50 (16)
Prednisone alone	37 (12)
Prednisone + others	24 (8)

TBI, total body irradiation; HLA, human leukocyte antigen; CNI, calcineurin inhibitor.

\*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.

prednisone at doses  $\geq 0.5$  mg/kg per day immediately before second-line treatment.

### Risk factors associated with treatment failure after second-line treatment

The FFS rate after second-line treatment was 56% (95% CI, 51%-62%) at 6 months (Figure 1). Cumulative incidence estimates of failure at 6 months by cause were 34% for treatment change (onset of third-line treatment), 7% for nonrelapse mortality during second-line treatment, and 3% for recurrent malignancy during second-line treatment. Cumulative incidence estimates of successful withdrawal of all immunosuppressive agents during second-line treatment were only 1% at 6 months and gradually reached 15% at 48 months.

In univariate analysis (Table 3), factors associated with increased risks of treatment failure included high-risk disease at transplantation; high-intensity conditioning with total-body irradiation compared with high-intensity conditioning without total-body irradiation; lower gastrointestinal involvement at second-line treatment;  $>3$  involved sites with chronic GVHD as compared with  $\leq 2$  involved sites; severe NIH global score of chronic GVHD compared with mild or moderate NIH global score; and thrombocytopenia, hyperbilirubinemia, and prednisone doses  $\geq 1$  mg/kg per day compared with no prednisone

**Table 2. GVHD characteristics at second-line treatment (N = 312)**

Characteristic	No. (%)
Median time from initial treatment to second-line treatment, range, mo	6.3 (0.2-61)
Reason for second-line treatment	
Progression	230 (74)
Lack of improvement	82 (26)
Sites involved	
Skin	226 (72)
Eyes	120 (38)
Mouth	221 (71)
Liver	91 (29)
Gastrointestinal tract	
Upper only	46 (15)
Any lower	47 (15)
Lung	35 (11)
Joint or fascia	73 (23)
Genital tract	31 (10)
Serosa	4 (1)
No. of sites involved	
1 or 2	127 (41)
3	94 (30)
4 or more	91 (29)
NIH global score	
Mild	26 (8)
Moderate	154 (50)
Severe	128 (42)
Not available	4
Platelet count	
$< 100\,000/\mu\text{L}$	54 (18)
$\geq 100\,000/\mu\text{L}$	250 (82)
Not available	8
Serum total bilirubin	
$> 2$ mg/dL	23 (7)
$\leq 2$ mg/dL	286 (93)
Not available	3
Prednisone-equivalent steroid dose immediately before second-line treatment	
None	46 (15)
$< 0.5$ mg/kg per day	140 (45)
$0.5$ - $1.0$ mg/kg per day	72 (23)
$\geq 1.0$ mg/kg per day	52 (17)
Not available	2
Second-line treatment	
Mycophenolate mofetil	95 (30)
Tacrolimus	63 (20)
Sirolimus	57 (18)
Extracorporeal photopheresis	23 (7)
Cyclosporine	17 (5)
Methotrexate	10 (3)
Other single agents	24 (8)
Multiple agents	23 (7)

immediately before second-line treatment. In multivariate analysis, 3 factors remained statistically significant: (1) high-risk disease at transplantation, (2) lower gastrointestinal involvement at second-line treatment, and (3) severe NIH global score at second-line treatment (Table 3). Thrombocytopenia was dropped early from the model because it correlated with lower gastrointestinal involvement and prednisone dose.

### Treatment failure rates according to risk groups

Treatment failure rates were analyzed for three risk groups categorized according to the total number of risk factors identified in the

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

Factor	Univariate		Multivariate	
	HR (95% CI)	P	HR (95% CI)	P
Time from transplantation to initial systemic treatment (per year)	0.73 (0.50-1.05)	.09		
Time from initial treatment to second-line treatment (per year)	0.88 (0.75-1.02)	.09		
Patient age at second-line treatment (per decade)	0.97 (0.90-1.04)	.40		
Female donor to male recipient	0.87 (0.66-1.16)	.36		
Disease risk at transplantation				
Low	1.00 (reference)		1.00 (reference)	
High	1.47 (1.12-1.93)	.005	1.49 (1.13-1.96)	.004
Conditioning regimen				
High intensity without TBI	1.00 (reference)			
High intensity with TBI	1.61 (1.17-2.23)	.004		
Reduced intensity	1.28 (0.94-1.74)	.11		
Graft source				
Bone marrow	1.00 (reference)			
Mobilized blood cells	0.83 (0.55-1.24)	.36		
Umbilical cord blood	1.36 (0.67-2.75)	.39		
Donor and HLA type				
HLA-matched related	1.00 (reference)			
HLA-matched unrelated	1.11 (0.83-1.49)	.49		
HLA-mismatched	1.31 (0.93-1.84)	.12		
Initial systemic treatment				
Prednisone + CNI	1.00 (reference)			
Prednisone + CNI + mycophenolate mofetil	1.28 (0.90-1.81)	.17		
Prednisone alone	0.83 (0.54-1.28)	.39		
Prednisone + others	1.49 (0.93-2.39)	.09		
Reason for second-line treatment				
Progression	1.00 (reference)			
Lack of improvement	0.95 (0.71-1.27)	.73		
Presence of involvement at second-line treatment				
Skin	1.22 (0.91-1.65)	.19		
Eyes	1.03 (0.79-1.34)	.86		
Mouth	1.19 (0.89-1.60)	.25		
Liver	1.27 (0.96-1.68)	.09		
Gastrointestinal tract				
No involvement	1.00 (reference)		1.00 (reference)	
Upper only	1.14 (0.79-1.64)	.48	1.15 (0.80-1.66)	.45
Any lower	1.62 (1.13-2.31)	.008	1.85 (1.29-2.65)	.0009
Lung	0.71 (0.46-1.09)	.12		
Joint or fascia	1.07 (0.79-1.44)	.66		
Genital tract	1.10 (0.71-1.69)	.67		
No. of sites involved at second-line treatment				
1 or 2	1.00 (reference)			
3	0.98 (0.72-1.35)	.91		
>3	1.36 (1.00-1.86)	.05		
NIH global score at second-line treatment				
Mild/moderate	1.00 (reference)		1.00 (reference)	
Severe	1.44 (1.11-1.87)	.006	1.49 (1.15-1.95)	.003
Thrombocytopenia at second-line treatment	1.42 (1.02-1.98)	.04		
Hyperbilirubinemia at second-line treatment	1.72 (1.10-2.70)	.02		
Prednisone-equivalent steroid dose immediately before second-line treatment				

**Table 3. (continued)**

Factor	Univariate		Multivariate	
	HR (95% CI)	P	HR (95% CI)	P
None	1.00 (reference)			
<0.5 mg/kg per day	0.86 (0.58-1.29)	.47		
0.5-1.0 mg/kg per day	0.92 (0.59-1.43)	.70		
≥1.0 mg/kg per day	1.61 (1.03-2.53)	.04		

CNI, calcineurin inhibitor; TBI, total-body irradiation.

\*Treatment failure was defined by the onset of third-line systemic treatment, nonrelapse mortality, or recurrent malignancy during second-line treatment.

multivariate model (Figure 2). This analysis excluded 4 patients who did not have information for NIH global severity score and 4 patients who had FFS at 6 months with information missing for steroid doses at 6 months. The low-risk group included 66 patients (22%) with no risk factor, the intermediate-risk group included 139 patients (46%) with 1 risk factor, and the high-risk group included 99 patients (33%) with 2 or 3 risk factors. As expected, failure rates were clearly stratified according to the three risk groups ( $P < .0001$ ).

**Cumulative incidence of successful withdrawal of immunosuppressive treatment according to steroid doses at 6 months**

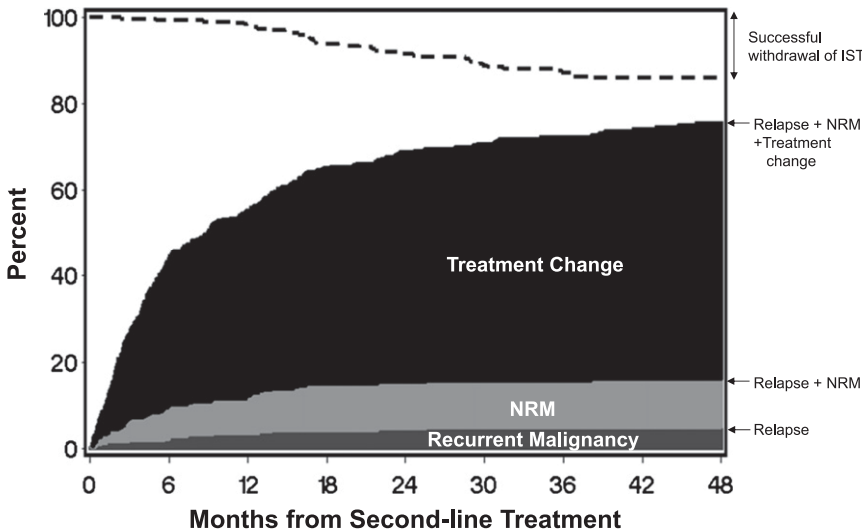
Steroid dose information at 6 months after second-line treatment was available in 170 of 174 patients who had FFS at 6 months. Among these patients, the median prednisone dose was 0.22 mg/kg per day (range, 0-2 mg/kg per day). Cumulative incidence estimates of successful withdrawal of all immunosuppressive agents during second-line treatment were compared according to various steroid dose limits at 6 months (Figure 3). The cumulative incidence of successful withdrawal was higher among patients taking prednisone-equivalent steroid doses  $\leq 0.3$ ,  $\leq 0.2$ , or  $\leq 0.1$  mg/kg per day than among those taking higher doses ( $P = .03$ ,  $.005$ , and  $.002$ , respectively). Cumulative incidence estimates of successful withdrawal at 48 months after second-line treatment were 29% for those with doses  $\leq 0.3$  mg/kg per day, 35% for those with doses  $\leq 0.2$  mg/kg per day, and 38% for those with doses  $\leq 0.1$  mg/kg per day.

**Shorter-term success rates at 6 months according to risk groups**

Table 4 summarizes FFS rates at 6 months according to risk groups, since FFS is the metric of interest for future clinical trials. Among all patients, the FFS rate was 0.56 at 6 months. The FFS rate was 0.67 in the low-risk group, 0.59 in the intermediate-risk group, and 0.44 in the high-risk group. When 6-month dose limits of prednisone  $\leq 0.3$ ,  $\leq 0.2$ , and  $\leq 0.1$  mg/kg per day were each incorporated as an additional criterion among all patients, success rates decreased progressively to 0.40, 0.25, and 0.14, respectively. Patterns were similar in each of the individual risk groups.

**Discussion**

This study offers several advances for future studies evaluating treatment of chronic GVHD. First, we show that FFS can serve as a meaningful end point for clinical trials. Second, we found that treatment change represented the predominant category of treatment failure, indicating that more efficient phase II trial design would be possible if treatment change were included as one type of failure in an FFS end point. Third, we identified 3 risk factors associated with



**Figure 1. Failure-free survival after second-line treatment of chronic GVHD.** 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 onset of third-line systemic treatment. The white area represents FFS. The dashed line represents cumulative incidence of successful withdrawal of all systemic immunosuppressive treatment (IST) during second-line treatment.

Relapse	3%	4%	4%	5%	5%	5%	5%	5%
NRM	7%	9%	11%	11%	11%	11%	11%	11%
Treatment change	34%	43%	50%	53%	55%	57%	58%	59%
FFS	56%	45%	35%	31%	29%	28%	26%	25%
Withdrawal of IST	1%	2%	6%	8%	11%	13%	14%	15%

failure and used them to stratify risk groups. Fourth, we found that lower steroid doses at 6 months were associated with higher rates of subsequent successful withdrawal of all immunosuppressive treatment during second-line treatment. Fifth, we report success rates accounting for both risk stratification at baseline and steroid dose limits at 6 months to make the end point more meaningful and to enhance the clinical benefit associated with the end point.

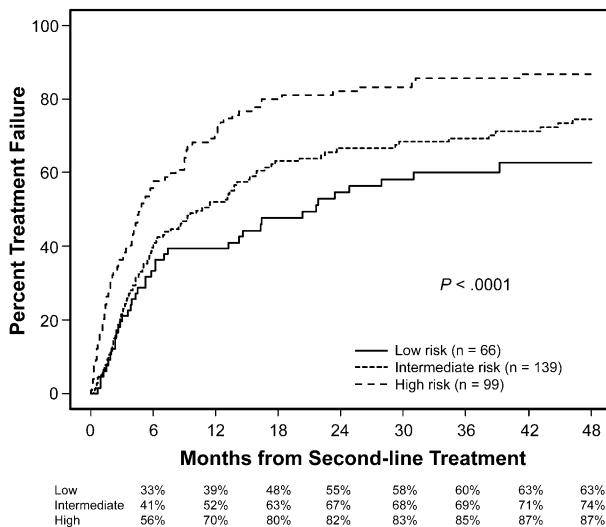
Our results showed that 34% of patients had second-line treatment failure at 6 months because of the initiation of third-line treatment, which represented the predominant cause of treatment failure; 7% had treatment failure because of nonrelapse mortality; and 3% had treatment failure because of recurrent malignancy. Thus, only 56%

had FFS at 6 months. Treatment failure was predicted by 3 clinical factors: high-risk disease at transplantation, lower gastrointestinal involvement at second-line treatment, and severe NIH global score at second-line treatment. Patients without any risk factors had a 67% FFS rate at 6 months, whereas those with 2 or more risk factors had an FFS rate of 44% at 6 months. High-risk disease at transplantation appeared to be associated with all 3 components of failure (data not shown). Gastrointestinal involvement and severe NIH global score have been associated with increased risk of mortality in previous studies.<sup>24-26</sup>

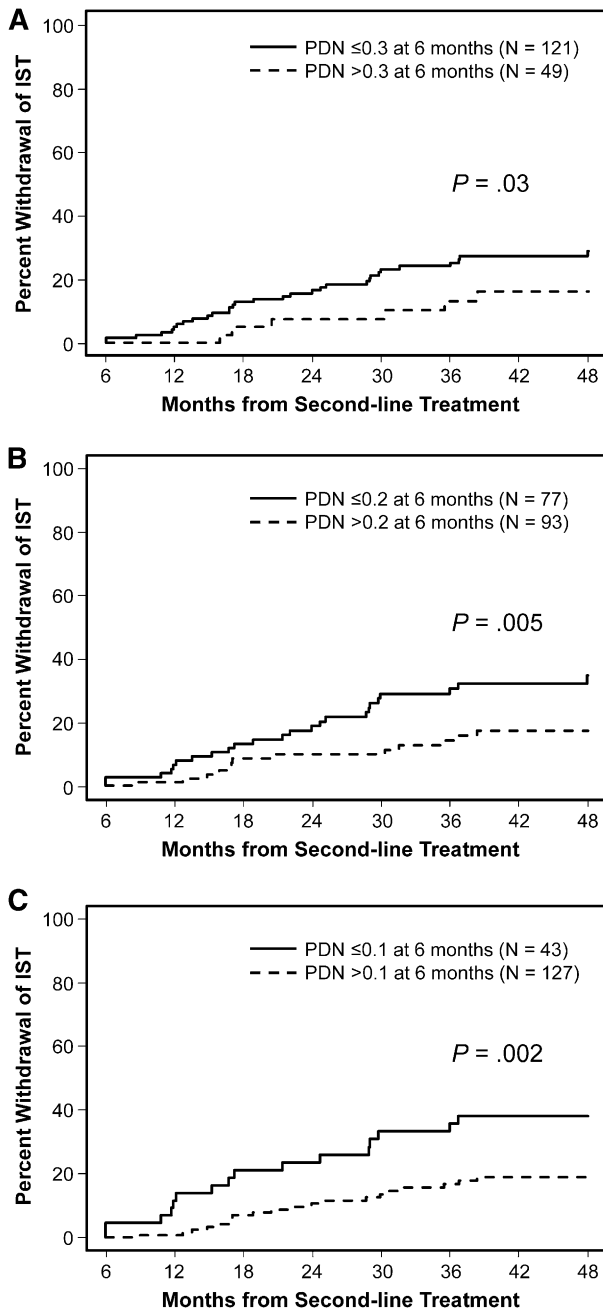
Describing the 3 causes of treatment failure helps to interpret the results of clinical studies. For example, careful interpretation would be required if the results showed an increased risk of nonrelapse mortality or recurrent malignancy despite a reduced risk of treatment change. For this purpose, results shown in Figure 1 provide a useful point of comparison for future studies.

Even for patients with FFS, the prolonged high-dose steroid exposure causes many adverse effects.<sup>27</sup> Therefore, it is important to consider steroid doses when the efficacy of second-line treatment is defined. The ability to control steroid-refractory chronic GVHD with reasonable doses of steroid is another important goal of second-line treatment. Our results showed that success end points incorporating a steroid dose limit at 6 months were associated with increased rates of subsequent withdrawal of all immunosuppressive treatment during second-line treatment, a long-term treatment goal that is usually not addressed in phase II trials. The choice of a dose limit to be used in defining success in future studies could depend on several factors, including patient characteristics, steroid doses administered at the onset of second-line treatment, and any expected steroid-sparing effects of the second-line agent. As in earlier studies,<sup>15,17,25,28,29</sup> we used withdrawal of all systemic immunosuppressive treatment without subsequent resumption for our analysis. As an alternative approach, current immunosuppressive treatment-free survival could also be analyzed to address the same question.<sup>30</sup>

The eligibility criteria for this study were carefully designed to represent the types of patients likely to be included in future clinical trials of second-line treatment of chronic GVHD, and all patients who met these criteria were included in the analysis. The minimum



**Figure 2. Cumulative incidence of treatment failure according to risk groups.** The low-risk group included patients with no risk factor, the intermediate-risk group included those with 1 risk factor, and the high-risk group included those with 2 or 3 risk factors. Risk factors included high-risk disease at transplantation, lower gastrointestinal involvement at second-line treatment, and severe NIH global score at second-line treatment.



**Figure 3. Successful withdrawal of systemic IST according to steroid doses at 6 months after second-line treatment.** (A)  $\leq 0.3$  mg/kg per day vs  $> 0.3$  mg/kg per day, (B)  $\leq 0.2$  mg/kg per day vs  $> 0.2$  mg/kg per day, and (C)  $\leq 0.1$  mg/kg per day vs  $> 0.1$  mg/kg per day. PDN, prednisone-equivalent steroid doses.

required duration of initial treatment before starting second-line treatment is somewhat shorter in our retrospective study than in previous prospective studies (2 weeks to 3 months).<sup>31-34</sup> Our detailed review of each patient demonstrated a well-justified rationale for second-line treatment, even within 1 month after initial treatment in some patients. For this reason, we propose that similar patients could be considered as candidates for future clinical trials. Therefore, the results in Table 4 could be used in the design and interpretation of future phase II trials of second-line systemic treatment of chronic GVHD. As an example, if a presumed study cohort contains 70% intermediate-risk patients and 30% high-risk patients, the probability of 6-month FFS with prednisone doses  $\leq 0.2$  mg/kg per day is

**Table 4. Shorter-term success rates at 6 months, according to risk groups at second-line treatment**

Risk group*	N (%)	Success rates at 6 mo			
		FFS	FFS + PDN $\leq 0.3$ mg/kg per day	FFS + PDN $\leq 0.2$ mg/kg per day	FFS + PDN $\leq 0.1$ mg/kg per day
All patients	304 (100)	0.56	0.40	0.25	0.14
Low	66 (22)	0.67	0.55	0.35	0.20
Intermediate	139 (46)	0.59	0.40	0.28	0.14
High	99 (33)	0.44	0.24	0.11	0.07

This analysis excluded 4 patients who did not have information for NIH global severity score and 4 patients who had FFS at 6 mo with information missing for steroid dose at 6 mo.

PDN, prednisone-equivalent steroid dose.

\*The low-risk group included patients with no risk factors, the intermediate-risk group included those with 1 risk factor, and the high-risk group included those with 2 or 3 risk factors. Risk factors included high-risk disease at transplantation, lower gastrointestinal involvement at second-line treatment, and severe NIH global score at second-line treatment.

expected to be 0.23 [estimated as  $(0.70 \times 0.28) + (0.30 \times 0.11)$ ]. If the expected success rate with the new treatment were 0.40 for this end point, enrollment of 46 patients would offer 81% statistical power with a 0.05 one-sided type I error, and a successful outcome in at least 16 patients would encourage further studies.

The use of FFS as a primary end point in clinical trials has some limitations, even if steroid dose limits are included as a component in the end point definition. First, results with this end point require careful interpretation in nonblinded trials, since treatment changes and the steroid taper schedule are controlled by providers. Nonetheless, treatment change implies that the second-line treatment did not provide the desired benefit and has always been considered as failure in evaluating the efficacy of a study intervention. To some extent, the objectivity of this end point could be improved by including a standardized guideline for tapering steroid doses and a standardized definition of progressive disease requiring treatment change in the protocol, although strict control of clinical management is difficult to enforce in practice, since both response and steroid toxicity may affect the speed of steroid taper. Second, this end point does not provide a direct measure of improved chronic GVHD activity or decreased symptom burden, although the use of a low steroid dose limit certainly serves as an indirect indication that GVHD is under good control in most cases. In the future, the FFS end point could further incorporate validated and standardized measures of response, reduced symptom burden, or improved quality of life as more comprehensive and direct indicators of clinical benefit.<sup>8,10,11</sup> Third, FFS at 6 months by itself does not address longer-term outcomes such as disability or withdrawal of immunosuppressive treatment. Although such end points are more appropriate for pivotal phase III studies, our results suggest that incorporation of an upper limit of the steroid dose in the shorter-term success end point of phase II trials could increase the probability that subsequent phase III studies would demonstrate differences in successful withdrawal of all immunosuppressive treatment as a longer-term end point.

Our results were derived from retrospective data at a single center and might differ from those at other centers because of variation in dosing and tapering schedules of immunosuppressive treatment. Additional studies are warranted to determine whether our results are representative of those from other centers, and prospective studies are needed to determine whether findings from retrospective studies hold true. Overall survival, relapse-free survival, and nonrelapse mortality have been recommended as long-term end points in clinical trials for treatment of chronic GVHD,<sup>5</sup> but failure events would occur much

earlier and more frequently if systemic treatment change were included as a component of the end point. FFS could be used as the basis for a clinically meaningful, efficient, and immediately applicable primary end point in the design and interpretation of phase II trials of second-line systemic treatment of chronic GVHD.

the Japan Society for the Promotion of Science Postdoctoral Fellowships for Research Abroad.

## Acknowledgments

The authors thank the patients who participated, members of the research staff, clinical staff, and referring physicians for their dedication and for their many years of service contributing to the long-term care of our patients after hematopoietic cell transplantation.

This work was supported by grants CA163438, CA18029, CA15704, and HL36444 from the National Institutes of Health, Department of Health and Human Services. Y.I. is a recipient of

## Authorship

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

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

Correspondence: Yoshihiro Inamoto, Fred Hutchinson Cancer Research Center, D5-290, PO Box 19024, Seattle, WA 98109-1024; e-mail: yinamoto@fhcrc.org.

## References

- Flowers ME, Storer B, Carpenter P, et al. Treatment change as a predictor of outcome among patients with classic chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2008;14(12):1380-1384.
- Wolff D, Schleuning M, von Harsdorf S, et al. Consensus Conference on Clinical Practice in Chronic GVHD: Second-Line Treatment of Chronic Graft-versus-Host Disease. *Biol Blood Marrow Transplant*. 2011;17(1):1-17.
- Lee SJ, Vogelsang G, Gilman A, et al. A survey of diagnosis, management, and grading of chronic GVHD. *Biol Blood Marrow Transplant*. 2002;8(1):32-39.
- Lee SJ, Joffe S, Artz AS, et al. Individual physician practice variation in hematopoietic cell transplantation. *J Clin Oncol*. 2008;26(13):2162-2170.
- Martin PJ, Weisdorf D, Przepiorka D, et al. Design of Clinical Trials Working Group. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: VI. Design of Clinical Trials Working Group report. *Biol Blood Marrow Transplant*. 2006;12(5):491-505.
- Martin PJ, Inamoto Y, Carpenter PA, et al. Treatment of chronic graft-versus-host disease: Past, present and future. *Korean J Hematol*. 2011;46(3):153-163.
- Inamoto Y, Flowers ME. Treatment of chronic graft-versus-host disease in 2011. *Curr Opin Hematol*. 2011;18(6):414-420.
- 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.
- 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.
- Palmer JM, Lee SJ, Chai X, et al. Poor agreement between clinician response ratings and calculated response measures in patients with chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2012;18(11):1649-1655.
- 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.
- Kim SJ, Lee JW, Jung CW, et al. Weekly rituximab followed by monthly rituximab treatment for steroid-refractory chronic graft-versus-host disease: results from a prospective, multicenter, phase II study. *Haematologica*. 2010;95(11):1935-1942.
- 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.
- Chien JW, Duncan S, Williams KM, et al. 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.
- 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.
- 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.
- 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.
- Gooley TA, Leisenring W, Crowley J, et al. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Stat Med*. 1999;18(6):695-706.
- Martin PJ, Carpenter PA, Sanders JE, et al. Diagnosis and clinical management of chronic graft-versus-host disease. *Int J Hematol*. 2004;79(3):221-228.
- Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2003;9(4):215-233.
- Smitten AL, Choi HK, Hochberg MC, et al. The risk of hospitalized infection in patients with rheumatoid arthritis. *J Rheumatol*. 2008;35(3):387-393.
- 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.
- 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.
- Lee SJ, Klein JP, Barrett AJ, et al. Severity of chronic graft-versus-host disease: association with treatment-related mortality and relapse. *Blood*. 2002;100(2):406-414.
- Cho BS, Min CK, Eom KS, et al. Feasibility of NIH consensus criteria for chronic graft-versus-host disease. *Leukemia*. 2009;23(1):78-84.
- Arai S, Jagasia M, Storer B, et al. Global and organ-specific chronic graft-versus-host disease severity according to the 2005 NIH Consensus Criteria. *Blood*. 2011;118(15):4242-4249.
- Khera N, Storer B, Flowers ME, et al. Nonmalignant late effects and compromised functional status in survivors of hematopoietic cell transplantation. *J Clin Oncol*. 2012;30(1):71-77.
- Stewart BL, Storer B, Storek J, et al. Duration of immunosuppressive treatment for chronic graft-versus-host disease. *Blood*. 2004;104(12):3501-3506.
- 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.
- 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.
- Courel DR, Saliba R, Escalón MP, et al. Sirolimus in combination with tacrolimus and corticosteroids for the treatment of resistant chronic graft-versus-host disease. *Br J Haematol*. 2005;130(3):409-417.
- Cutler C, Miklos D, Kim HT, et al. Rituximab for steroid-refractory chronic graft-versus-host disease. *Blood*. 2006;108(2):756-762.
- Jacobsohn DA, Chen AR, Zahurak M, et al. Phase II study of pentostatin in patients with corticosteroid-refractory chronic graft-versus-host disease. *J Clin Oncol*. 2007;25(27):4255-4261.
- Chen GL, Arai S, Flowers ME, et al. A phase 1 study of imatinib for corticosteroid-dependent/refractory chronic graft-versus-host disease: response does not correlate with anti-PDGFRα antibodies. *Blood*. 2011;118(15):4070-4078.