

Failure-free survival in a prospective cohort of patients with chronic graft-versus-host disease

Jeanne Palmer,¹ Xiaoyu Chai,² Paul J Martin,² Daniel Weisdorf,³ Yoshihiro Inamoto,^{2,4} Joseph Pidala,⁵ Madan Jagasia,⁶ Steven Pavletic,⁷ Corey Cutler,⁸ Georgia Vogelsang,⁹ Sally Arai,¹⁰ Mary ED Flowers,² and Stephanie J Lee²

¹Division of Hematology/Oncology, Mayo Clinic Phoenix, AZ, USA; ²Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; ³Blood and Marrow Transplant Program, University of Minnesota, Minneapolis, MN, USA; ⁴Division of Hematopoietic Stem Cell Transplantation National Cancer Center Hospital, Tokyo, Japan; ⁵Blood and Marrow Transplantation, Moffitt Cancer Center, Tampa, FL, USA; ⁶Hematology and Stem Cell Transplant Program, Vanderbilt University Medical Center, Nashville, TN, USA; ⁷National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; ⁸Hematologic Malignancies, Dana-Farber Cancer Institute, Boston, MA, USA; ⁹Division of Oncology, Johns Hopkins Hospital, Baltimore, MD, USA; and ¹⁰Division of Blood and Marrow Transplantation, Stanford University Medical Center, CA, USA

ABSTRACT

Failure-free survival, defined as the absence of relapse, non-relapse mortality or addition of another systemic therapy, has been proposed as a potential endpoint for clinical trials, but its use has only been reported for single-center studies. We measured failure-free survival in a prospective observational cohort of patients (n=575) with both newly diagnosed and existing chronic graft-versus-host disease from nine centers. Failure was observed in 389 (68%) patients during the observation period. The median follow up of all patients was 30.9 months, and the median failure-free survival was 9.8 months (63% at 6 months, 45% at 1 year, and 29% at 2 years). Of the variables measured at enrollment, ten were associated with shorter failure-free survival: higher National Institutes of Health 0-3 skin score, higher National Institutes of Health 0-3 gastrointestinal score, worse range of motion summary score, lower forced vital capacity (%), bronchiolitis obliterans syndrome, worse quality of life, moderate to severe hepatic dysfunction, absence of treatment for gastric acid, female donor for male recipient, and prior grade II-IV acute graft-versus-host disease. Addition of a new systemic treatment, the major cause of failure, was associated with an increased risk of subsequent non-relapse mortality (hazard ratio=2.06, 95% confidence interval: 1.29-3.32; $P<0.003$) and decreased survival (hazard ratio=1.51, 95% confidence interval: 1.04-2.18; $P<0.03$). These results show that fewer than half of patients on systemic treatment will be failure-free survivors at 1 year, and fewer than a third will reach 2 years without experiencing failure. Better treatments are needed for chronic graft-versus-host disease. *Clinicaltrials.gov identifier: NCT00637689.*

Introduction

Chronic graft-versus-host disease (GVHD) is a significant cause of morbidity and mortality in survivors of allogeneic hematopoietic cell transplantation.^{1,8} Recently, Inamoto *et al.* proposed a novel measure of response: failure-free survival (FFS), defined as the absence of an additional systemic therapy, relapse or non-relapse mortality.^{9,10} In a single center, the FFS rates after 12 months of treatment were 54% in newly diagnosed patients and 45% in those on second-line treatment. Predictive factors for shorter FFS in patients receiving initial treatment were onset of initial systemic treatment within the first year after transplant, patient's age >60 years, severe gastrointestinal, liver or lung involvement and Karnofsky performance score <80%. Risk factors in patients receiving second-line treatment were high-risk disease, lower gastrointestinal tract involvement and severe National Institutes of Health global score of chronic GVHD (NIH global score). In the current study, we examined FFS rates in a multicenter and heterogeneous cohort including incident and prevalent cases of chronic GVHD, to identify other clinical factors predicting FFS, and to determine whether addition of a new therapy is associated with subsequent mortality.

Methods

Patients

A cohort of hematopoietic cell transplantation recipients affected by chronic GVHD was enrolled in a multicenter observational study (NCT00637689).¹¹ Full details regarding the study are described elsewhere.¹² The protocol was approved by the Institutional Review Board at each site, and all subjects provided written informed consent.

Definitions

Chronic GVHD was defined by NIH consensus criteria.¹³ Failure was defined by relapse, non-relapse mortality or addition of a new immunosuppressive medication intended for systemic treatment of chronic GVHD.¹⁰ Please see the *Online Supplement* for details. Treatment failure was determined by two separate reviewers (JP and SJL) independently and discrepancies were resolved by discussion.

Potential predictors

To identify enrollment variables associated with FFS, univariable analyses were performed on all available information. The complete list is provided in *Online Supplementary Table S1*.

©2015 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2014.117283

The online version of this article has a *Supplementary Appendix*.

Manuscript received on September 22, 2014. Manuscript accepted February 20, 2015.

Correspondence: palmer.jeanne@mayo.edu

Statistical analysis

Failure-free survival was estimated from the time of study enrollment. Cumulative incidence estimates of relapse, non-relapse mortality and addition of a new therapy as causes of failure were derived, treating each event as a competing risk for the other two.¹⁴

Cox regression models were used to identify risk factors for failure, using sequential selection processes within each organ because of the large number of potential predictors, many of which were correlated.

We also examined the association between adding a new systemic therapy and subsequent survival outcomes. Cox regression analysis was used to model overall mortality and non-relapse mortality from the time of enrollment, with treatment change included as a time-varying covariate. We did not analyze the association of relapse with survival, since this relationship is well established.

Statistical analyses were performed using SAS/STAT software, version 9.3 (SAS Institute, Inc., Cary, NC, USA) and R version 2.15.2 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Patients' and graft-versus-host disease characteristics

As of January 31, 2013, 575 patients were enrolled in the prospective observational study (Table 1). There were 1,856 follow-up visits, for a total of 2,431 visits. The cohort included 342 (59%) incident cases and 233 (41%) prevalent cases. In prevalent cases, the median time between chronic GVHD diagnosis and enrollment was 10.3 months (interquartile range, 6.3-16.8 months). The majority of the patients had overlap chronic GVHD (n=477, 83%), which was determined retrospectively by chart review for prevalent cases. The median time of follow up for survivors was 30.9 months (range, 0.9-65.8 months). Of the 53 patients with "mild chronic GVHD or less," 51 had mild chronic GVHD while two technically had less than mild, meaning that they scored a 0 in all organ scores, even though they met the NIH diagnostic criteria for chronic GVHD.

FFS for the entire cohort was 63% at 6 months, 45% at 1 year, and 29% at 2 years (Figure 1). Among the 575 patients, 389 experienced failure due to addition of systemic therapy (n=300, 77%), relapse (n=54, 14%) and non-relapse mortality (n=35, 9%) (Figure 2). The median FFS for the entire study cohort was 9.8 months [95% confidence interval (95% CI): 9.0-11.7] (Table 2). When analyzed separately, incident cases tended to have a shorter median FFS than that of prevalent cases (9.2 months *versus* 11.6 months, respectively; $P=0.18$). Severity of the NIH global score at enrollment was associated with FFS. Patients with severe chronic GVHD had a median FFS of 8.5 months (95% CI: 5.6-9.6), whereas those with moderate or mild GVHD had a median FFS of 11.7 months (95% CI: 9.3-16.8) and 14.9 months (95% CI: 8.3-23.3), respectively ($P<0.001$, Table 2 and Figure 1). When relapse was excluded from the definition of failure, the median FFS was approximately 1 month longer (11.0 months, 95% CI: 9.3-14.8).

Risk factors associated with failure-free survival

All tested variables are reported in *Online Supplementary Table S1*. *Online Supplementary Table S2* shows the results of univariable and multivariable analyses for each organ system. Table 3 shows the factors that were statistically

associated with FFS at $P\leq 0.05$ in multivariable analysis including all variables identified in the organ-specific analysis. Two variables preceded the diagnosis of chronic GVHD: female donor for male recipient and history of grade II-IV acute GVHD and were associated with a higher risk of failure. Four factors directly measured GVHD severity [higher NIH 0-3 skin score, higher NIH 0-3 gastrointestinal score, worse (lower) range of motion joint scores, presence of bronchiolitis obliterans syndrome] and were associated with a higher risk of failure. Higher scores of one laboratory test (forced vital capacity) and one measure of quality of life, the Functional Assessment of Cancer

Table 1. Patients' characteristics.

Characteristic	Category	N.	Count (%)
Study site	Fred Hutchinson Cancer Research Center	575	253 (44%)
	University of Minnesota		61 (11%)
	Dana-Faber Cancer institute		65 (11%)
	Stanford University Medical Center		74 (13%)
	Vanderbilt University Medical Center		48 (8%)
	Medical College of Wisconsin		23 (4%)
	Washington University Medical Center		4 (1%)
	Moffitt Cancer Center		39 (7%)
	Memorial Sloan-Kettering Cancer Center		8 (1%)
Case type	Incident	575	342 (59%)
	Prevalent		233 (41%)
Patients' age at registration (years), median (range)		575	52 (19-79)
Patients' gender	Female	575	242 (42%)
	Male		333 (58%)
Diagnosis	Acute myeloid leukemia	575	193 (34%)
	Acute lymphoblastic leukemia		62 (11%)
	Chronic myeloid leukemia		30 (5%)
	Chronic lymphocytic leukemia		46 (8%)
	Myelodysplastic syndrome		89 (15%)
	Non-Hodgkin lymphoma		85 (15%)
	Hodgkin lymphoma		17 (3%)
	Multiple myeloma		29 (5%)
	Aplastic anemia		6 (1%)
	Other		18 (3%)
Disease status	Early	571	184 (32%)
	Intermediate		248 (43%)
	Advanced		139 (24%)
Transplant source	Bone marrow	575	40 (7%)
	Cord blood		28 (5%)
	Peripheral blood		507 (88%)
Transplant type	Myeloablative	571	297 (52%)
	Non-myeloablative		274 (48%)
Donor-patient CMV status	Patient and donor CMV both negative	569	192 (34%)
	Patient or donor CMV positive		377 (66%)
Donor-patient gender combination	Female into male	569	167 (29%)
			402 (71%)
Donor match	Matched related	573	238 (42%)
	Matched unrelated		244 (43%)
	Mismatched		91 (16%)
Prior grade II-IV acute GVHD	Yes	575	311 (54%)
	No		264 (46%)
NIH chronic GVHD global severity score	Mild or less	575	53 (9%)
	Moderate		302 (53%)
	Severe		220 (38%)

CMV: cytomegalovirus; GVHD: graft-versus-host disease; NIH: National Institutes of Health.

Table 2. Description of FFS.

Characteristics	Number failed / total	Median FFS months (95% CI)	6-month FFS (95% CI)	1-year FFS (95% CI)	2-year FFS (95% CI)
Overall	389/575	9.8 (9.0-11.7)	63% (58%-66%)	45% (41%-49%)	29% (25%-33%)
Case type					
Incident	232/342	9.2 (7.8-11.0)	59% (54%-65%)	43% (38%-49%)	29% (24%-35%)
Prevalent	157/233	11.6 (9.1-15.1)	67% (61%-73%)	48% (41%-55%)	29% (22%-35%)
NIH global severity					
Mild or less	36/53	14.9 (8.3-23.3)	69% (55%-80%)	53% (38%-65%)	32% (18%-47%)
Moderate	186/302	11.7 (9.3-16.8)	66% (60%-71%)	50% (44%-55%)	37% (31%-43%)
Severe	167/220	8.5 (5.6-9.6)	56% (49%-63%)	38% (31%-44%)	17% (11%-23%)
Chronic GVHD type					
Classic	61/98	11.5 (8.9-15.4)	68% (58%-77%)	48% (37%-57%)	29% (19%-40%)
Overlap	328/477	9.7 (8.6-11.6)	61% (57%-66%)	45% (40%-49%)	29% (24%-33%)

NIH: National Institutes of Health; GVHD: graft-versus-host disease.

Therapy Bone Marrow Transplant module, Trial Outcome Index (FACT BMT TOI) were associated with better functioning and a lower risk of failure. Two comorbidities were also associated with FFS: the presence of moderate to severe hepatic dysfunction, defined as liver cirrhosis, total serum bilirubin concentration >1.5 times the upper limit of normal or transaminase concentration >2.5 times the upper limit of normal, was associated with shorter FFS, while a co-morbidity of peptic ulcer disease, hiatal hernia, reflux disease, or treatment with an acid-reducing agent was associated with longer FFS. Notably, incident *versus* prevalent case status at time of enrollment was not associated with FFS. No statistically significant interactions were detected between significant variables in the multivariate analysis.

Addition of systemic immunosuppressive medication

Addition of new systemic treatment for chronic GVHD accounted for most of the failure events. A time-varying Cox regression analysis showed that addition of a new medication was associated with increased risks of overall mortality [hazard ratio (HR)=1.51, 95% CI: 1.04-2.18; $P=0.03$] and non-relapse mortality (HR=2.06, 95% CI: 1.29-3.32; $P=0.003$).

Discussion

In this study, we used a prospectively studied multicenter cohort of patients with chronic GVHD to explore the parameters of FFS and to characterize the factors that predict failure. We identified ten factors, including pre-transplant variables (female donor for male recipient and history of grade II-IV acute GVHD), laboratory values (a comorbidity of hepatic dysfunction defined by liver function tests and lower forced vital capacity), and clinical findings (higher NIH 0-3 skin score, high NIH 0-3 gastrointestinal score, worse range of motion score, absence of gastroin-

Table 3. Multivariable analysis of factors associated with FFS.

Variable measured	HR	95% CI	P value
NIH 0-3 skin score	1.22	1.11-1.34	<0.001
NIH 0-3 gastrointestinal score	1.20	1.01-1.41	0.04
Range of motion summary score*	0.89	0.84-0.95	0.001
Forced vital capacity %*	0.99	0.98-0.997	0.008
Bronchiolitis obliterans syndrome	1.93	1.22-2.92	0.007
FACT BMT TOI*	0.99	0.986-1.00	0.05
Comorbidity peptic ulcer/hernia/reflux	0.73	0.58-0.93	0.01
Comorbidity moderate/severe hepatic	1.78	1.16-2.62	0.01
Female donor for male recipient	1.30	1.04-1.61	0.02
Prior grade II-IV acute GVHD	1.34	1.09-1.65	0.006

*Lower scores represent greater dysfunction. HR: hazard ratio; NIH: National Institutes of Health; FACT BMT TOI: Functional Assessment of Cancer Therapies, Bone Marrow Transplant Subscale, Trial Outcome Index; GVHD: graft-versus-host disease.

testinal co-morbidity of peptic ulcer disease, hiatal hernia, or reflux disease, and presence of bronchiolitis obliterans syndrome) that predict shorter FFS. Shorter FFS was also associated with worse quality of life (lower score) on the FACT BMT TOI. Of note both forced vital capacity and the FACT BMT TOI had hazard ratios close to 1 because these scales are continuous with a large dynamic range. In our cohort, FFS was 63% at 6 months, 45% at 1 year, and 29% at 2 years, proportions which are similar to those observed in previous studies.^{9,10} We also demonstrated that addition of a new treatment was associated with higher non-relapse and overall mortality, which is in agreement with previous studies.^{15,16}

Inamoto *et al.* published two analyses of FFS in patients who received first-line⁹ or second-line¹⁰ therapy for chronic GVHD at a single institution. The median duration of

FFS and the rates of FFS at 6 months and 2 years in the current study are similar to their results^{9,10} despite the differences in cohort definition. Failure of first-line therapy was associated with grade 3 gastrointestinal, liver or lung involvement, Karnofsky performance score <80% at enrollment, age >60 years, and shorter interval between transplant and onset of chronic GVHD.⁹ Failure of second-line therapy was associated with higher disease risk at the time of the transplant, lower gastrointestinal tract involvement and severe chronic GVHD.¹⁰ The factors that predicted FFS were similar in our study, including lower gastrointestinal tract involvement, and lung involvement. Although liver involvement did not predict outcome, hepatic comorbidity did predict FFS. Both are defined by liver function test abnormalities, although the thresholds differ. Global chronic GVHD severity was significant in our univariate analysis but was not included in the final multivariate analysis due to the overlap with organ-specific information. Finally, Karnofsky performance score and age were not found to be significantly associated with outcome in our analysis. There were some shared patients in the two cohorts. Forty-eight (8%) of the patients in our study were also included in Inamoto's secondary analysis cohort,¹⁰ although these patients were analyzed at different times in the course of their chronic GVHD. There was a higher degree of overlap with the FFS analysis performed for first-line therapy,⁹ in which 167 (29%) of patients overlapped with our cohort. However, when patients shared with the Inamoto's study were excluded, the results were unchanged (*data not shown*).

Our study confirms that the severity of skin, joint, gastrointestinal tract, lung and liver GVHD predicts FFS, which is driven primarily by treatment changes. These characteristics have also been reported to be important in other analyses evaluating predictors of survival. For example, lower gastrointestinal tract and liver involvement

have been shown to be associated with an inferior survival when considered alone¹⁷ or as part of the overlap subtype.¹⁸ Liver involvement, low Karnofsky performance score and older age were associated with inferior survival in a study of a separate cohort of patients¹⁹ in which other specific manifestations of chronic GVHD were not analyzed. Higher Lee skin symptom score and NIH 0-3 skin score have been associated with increased non-relapse mortality.²⁰ However, in other studies, skin manifestations have not been associated with non-relapse mortality or overall survival.^{12,21} The clinical manifestations of bronchiolitis obliterans syndrome and a decreased forced vital capacity are associated with shorter FFS, which is not surprising, as the poor prognosis of pulmonary GVHD has been noted in many studies.^{22,23} The global severity of GVHD was significant in our univariate analysis but was not included in the final multivariate analysis due to the overlap with organ-specific information.

It should be noted that the involvement of some organs – eyes, mouth and genitals – was not associated with FFS. This finding is not unexpected as these organs are usually treated with topical therapy. However, it is notable that factors historically associated with a worse outcome in chronic GVHD, such as low platelet counts, lower Karnofsky performance score, and progressive onset were significant in univariate analysis, but not in the organ-specific or overall multivariate analyses. The reason for these discrepant findings may be due to differences in endpoints (6-month FFS *versus* non-relapse mortality or survival) or analysis approach. In general, we were able to consider many more potentially predictive variables than prior analyses because of our detailed data collection. Finally, disease status at transplant and age were not found to be significantly associated with FFS in our analysis.

The strengths of our study include the large number of patients evaluated, the prospective collection of detailed

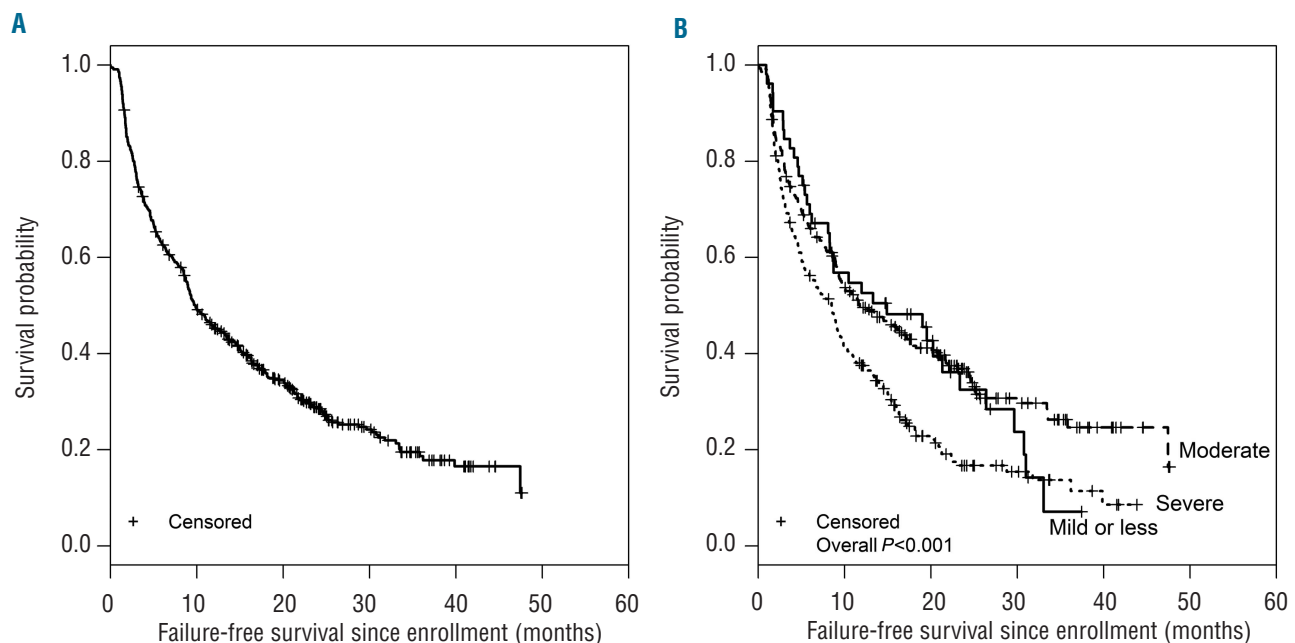


Figure 1. FFS of (A) the entire cohort and (B) separately by NIH global severity of GVHD at enrollment.

data from multiple centers, and the comprehensive consideration of variables in regression analyses. Two limitations should, however, be highlighted. Enrollment in the cohort did not depend on treatment status. The cohort included patients with minimal or longer-term stable immunosuppressive regimens and patients with much more intensive treatment or several prior lines of therapy, resulting in a heterogeneous population of patients with different risks of failure. We did not have accurate information regarding either the number of lines of therapy or the intensity of the therapies. Regardless, we did not detect significant differences according to several classification systems, including incident *versus* prevalent cases or classic *versus* overlap subtypes of GVHD. Second, the reasons for treatment changes were not recorded. For example, we could not determine whether a given change from tacrolimus to mycophenolate mofetil was related to toxicity such as thrombotic microangiopathy or neurotoxicity, or whether the change was prompted by ineffectiveness of the first medication. This has been a criticism of the FFS endpoint for clinical trials. Nevertheless, patients in clinical trials who have systemic treatment added for any reason are usually considered treatment failures, and the absence of new treatment serves as a minimal requirement for success when the primary endpoint is assessed.

The heterogeneity of our population may be considered both a strength and a limitation. Different transplant centers likely have different practices regarding steroid tapering which may affect the failure rates. Specific steroid doses were not available unless they were given as high-dose regimens. Additionally, the clinical trial portfolio varies from center to center, and one reason to start new therapy may be the option to provide a novel therapy on a clinical trial. Although these are potential confounders, they also increase the generalizability of our findings as they are more likely to be representative of the variety of patients who would be enrolled in a multicenter clinical trial. It is notable that our findings are similar to those of Inamoto *et al.*^{9,10} in a more homogeneous population that had minimal overlap.

There were two findings that were somewhat unexpected. First, the presence of a co-morbidity of peptic ulcer disease, hiatal hernia or gastric acid reflux disease (n=169) was associated with a prolonged FFS. As patients who were on acid-reducing medications were also considered to have this co-morbidity, it may reflect higher doses or more prolonged treatment with glucocorticoids. We were unable to analyze steroid doses from the available data. There did not appear to be an association between peptic ulcer comorbidity and severity of NIH global score (*data not shown*). Second, overlap syndrome was not associated with shorter FFS. Previous studies found that overlap syndrome was associated with decreased overall survival. However, half the failures occurred within the first year and were primarily due to treatment change, not death, potentially explaining why our current results differ from those of prior studies.

One may question whether relapse should be considered a failure of chronic GVHD therapy, as the risk of relapse depends more on the malignant disease risk rather than any features of chronic GVHD. In agreement with Inamoto *et al.*,^{9,10} we included this component in the endpoint, because potent immunosuppression that could con-

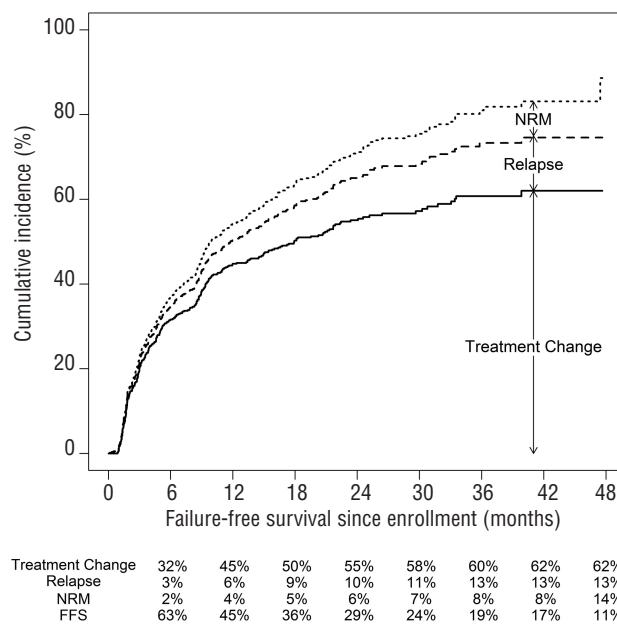


Figure 2. Cumulative incidence of treatment change, treatment change plus relapse, and treatment change plus relapse plus non-relapse mortality (NRM).

trol GVHD could increase the risk of relapse. When we reanalyzed the data excluding relapse from the definition of failure, the median FFS was only 1 month longer (*data not shown*). Additionally, disease relapse contributed to only a minority of the failures, and disease risk did not have a statistically significant association with FFS in the univariate analysis, suggesting that disease relapse is not a major driver of the duration of FFS.

In summary, FFS correlates with subsequent non-relapse mortality and survival. Absence of these failure events should be recognized as the minimal definition of success for an investigational agent. Our results highlight the poor outcomes in patients with chronic GVHD and the unsatisfactory ability of currently available therapies to control the disease adequately. By 6 months after enrollment into our study, a third of patients had already been started on new systemic agents, relapsed or died. By 12 months after enrollment, more than half had failed in one of these ways, and by 2 years after enrollment only 30% of patients had not relapsed, died, or started another treatment. These results clearly illustrate the need for new, more effective, and less toxic therapies for chronic GVHD.

Funding

The Chronic GVHD Consortium (U54 CA163438) is a part of the National Institutes of Health Rare Disease Clinical Research Network, supported through a collaboration between the Office of Rare Diseases Research, NCATS, and the National Cancer Institute. This work was also supported by CA118953

Authorship and Disclosures

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

- Socie G, Ritz J, Martin PJ. Current challenges in chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2010;16(1):S146-S151.
- Socie G, Salooja N, Cohen A, et al. Nonmalignant late effects after allogeneic stem cell transplantation. *Blood.* 2003;101(9):3373-3385.
- Socie G, Stone JV, Wingard JR, et al. Long-term survival and late deaths after allogeneic bone marrow transplantation. *N Engl J Med.* 1999;341(1):14-21.
- Lee SJ, Kim HT, Ho VT, et al. Quality of life associated with acute and chronic graft-versus-host disease. *Bone Marrow Transplant.* 2006;38(4):305-310.
- Lee SJ, Vogelsang G, Flowers MED. Chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2003;9(4):215-233.
- Fraser CJ, Bhatia S, Ness K, et al. Impact of chronic graft-versus-host disease on the health status of hematopoietic cell transplantation survivors: a report from the Bone Marrow Transplant Survivor Study. *Blood.* 2006;108(8):2867-2873.
- 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.
- Socie G, Schmoor C, Bethge WA, et al. 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.
- Inamoto Y, Flowers MED, Sandmaier BM, et al. Failure-free survival after initial systemic treatment of chronic graft-versus host disease. *Blood.* 2014;124(8):1363-1371.
- 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.
- Chronic GVHD Consortium. Rationale and design of the chronic GVHD cohort study: improving outcomes assessment in chronic GVHD. *Biol Blood Marrow Transplant.* 2011;17(8):1114-1120.
- Inamoto Y, Storer BE, Petersdorf EW, et al. Incidence, risk factors, and outcomes of sclerosis in patients with chronic graft-versus-host disease. *Blood.* 2013;121(25):5098-5103.
- 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.
- 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.
- Flowers MED, 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.
- Kim DH, Sohn SK, Baek JH, et al. Time to first flare-up episode of GVHD can stratify patients according to their prognosis during clinical course of progressive- or quiescent-type chronic GVHD. *Bone Marrow Transplant.* 2007;40(8):779-784.
- Pidala J, Chai X, Kurland BF, et al. Analysis of gastrointestinal and hepatic chronic graft-versus-host disease manifestations on major outcomes: a chronic graft-versus-host disease consortium study. *Biol Blood Marrow Transplant.* 2013;19(5):784-791.
- Pidala J, Vogelsang G, Martin P, et al. Overlap subtype of chronic graft vs. host disease is associated with adverse prognosis, functional impairment, and inferior patient reported outcomes: a Chronic Graft vs. Host Disease Consortium study. *Haematologica.* 2011;96(11):1678-1684.
- 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.
- Jacobsohn DA, Kurland BF, Pidala J, et al. Correlation between NIH composite skin score, patient-reported skin score, and outcome: results from the Chronic GVHD Consortium. *Blood.* 2012;120(13):2545-2552.
- Baird K, Steinberg SM, Grkovic L, et al. National Institutes of Health chronic graft-versus-host disease staging in severely affected patients: organ and global scoring correlate with established indicators of disease severity and prognosis. *Biol Blood Marrow Transplant.* 2013;19(4):632-639.
- Williams KM, Chien JW, Gladwin MT, Pavletic SZ. Bronchiolitis obliterans after allogeneic hematopoietic stem cell transplantation. *JAMA.* 2009;302(3):306-314.
- 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):S106-S114.