Incidence and Outcome of Chronic Graft-versus-Host Disease Using National Institutes of Health Consensus Criteria

Madan Jagasia,[1] Jennifer Giglia,[1] Wichai Chinratanalab,[1,2] Sheri Dixon,[1] Heidi Chen,[3] Haydar Frangoul,[4] Brian Engelhardt,[1] Stacey Goodman,[1,2] John Greer,[1] Adetola Kassim,[1] David Morgan,[1] Katherine Ruffner,[1,2] Friedrich Schuening[1]

[1]Division of Hematology-Oncology, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee; [2]Veterans Administration Hospital–Nashville, Nashville, Tennessee; [3]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee; and [4]Division of Hematology-Oncology, Department of Pediatrics, Vanderbilt University Medical Center, Nashville, Tennessee

Correspondence and reprint requests: Madan Jagasia, MBBS, MS, 2665 TVC, Division of Hematology-Oncology, Department of Medicine, Vanderbilt Ingram Cancer Center, Department of Internal Medicine, Vanderbilt University Medical Center, Nashville, TN 37232-5505 (e-mail: madan.jagasia@vanderbilt.edu).

ABSTRACT

Chronic graft-versus-host disease (cGVHD), a common complication after stem cell transplant (SCT), has an impact on morbidity and survival. Previous classification of cGVHD has not been reproducible or prognostic for nonrelapse mortality (NRM). Recently the National Institutes of Health (NIH) consensus criteria were proposed, but the ability of this classification to predict outcome of various subtypes of cGVHD is unknown. Patients (N = 110) undergoing an SCT for a hematologic malignancy and surviving until day 100 posttransplant from 2001 to 2003 were studied. The overall survival (OS) using a landmark analysis at day 100 was 44% versus 66% (no GVHD vs. GVHD, P = .026). The OS of patients with various types of GVHD as proposed by the NIH criteria were significantly different (P < .0001). In a univariate analyses, this was more apparent when patients with any acute features of GVHD were compared to classic cGVHD (3-year OS 46% vs. 68%, P = .033). The 3-year NRM for the entire cohort was 21%, and was not affected by presence or absence of GVHD or subtypes of GVHD. In a multivariable analysis, extensive cGVHD (hazard ratio [HR] 0.35, P = .015) and having any acute feature of GVHD after day 100 (HR 3.36, P = .0144) were significant independent predictors of survival. The OS with different NIH subtypes of GVHD after day 100 from SCT varies, and is superior for patients with classic cGVHD.
© 2007 American Society for Blood and Marrow Transplantation

KEY WORDS

Chronic GVHD • Grading • NIH consensus criteria

INTRODUCTION

Graft-versus-host disease (GVHD) is a common complication occurring after allogeneic stem cell transplant (SCT) [1,2]. Traditionally, this has been classified as acute (onset within first 100 days of allogeneic SCT) and chronic (onset after 100 days of allogeneic SCT). Acute GVHD (aGVHD) presents as dermatitis, enteritis, and/or hepatitis, and is treated with short courses of high-dose steroids followed by a variable taper over several weeks [1]. Chronic GVHD (cGVHD) has protean manifestations that resemble a variety of autoimmune disorders. The most common classification still in use stratifies cGVHD as limited and extensive based on clinical and laboratory features [3,4]. The prognostic significance of this classification is unclear. Recently, the Seattle group published a revised classification, which stratifies GVHD into clinical limited and extensive, and is based on morphologic manifestations and extent of cGVHD. This classification allows for decision making regarding systemic immunosuppression for patients with clinical extensive GVHD [4]. Both these classifications for cGVHD still use the 100-day time point, after which

any GVHD is called cGVHD. The National Institutes of Health (NIH) published consensus criteria for diagnosis of cGVHD. Manifestations of GVHD are classified according to various organ systems and sites involved into distinctive, diagnostic, associated with, or common to both aGVHD and cGVHD. GVHD after day 100 can now be classified as acute (persistent, recurrent, or late acute), classic cGVHD (features of cGVHD only), or overlap GVHD (features of acute and classic chronic). Each target organ affected is scored for severity of involvement taking into account not only clinical or laboratory criteria, but also functionality. Using this composite score, a new global assessment of cGVHD severity has been proposed [5]. The goal of this classification is to establish a common platform for subclassifying GVHD. The impact of these revised criteria on the incidence and outcome of cGVHD remains unknown. It is important to validate these criteria, to try and identify a subset of patients who may be at a higher risk of mortality, so future clinical studies can target this patient population.

In this study, we have retrospectively reclassified GVHD present after day 100 using the NIH consensus criteria, and show the impact of the revised classification on the incidence and outcome of cGVHD. Our aim was to see if the presence of various subtypes of GVHD after day 100 as proposed by the NIH consensus criteria affects survival or nonrelapse mortality (NRM). The NIH consensus criteria have established a severity scale for patients with cGVHD. We have assessed the impact of severity in predicting survival or NRM among patients with cGVHD.

PATIENTS AND METHODS

All patients undergoing an allogeneic SCT for a hematologic malignancy and surviving until day 100 after transplant between January 1, 2001, and December 31, 2003, at Vanderbilt University Medical Center (adults and pediatrics) and the associated Veterans Administration Hospital were included in the initial cohort (196 patients). All patients underwent transplant on standard of care or institutional review board (IRB) approved protocols. All patients signed IRB-approved data consents. To study a homogenous cohort, patients with the following criteria were excluded: received 2 allogeneic SCT, received donor lymphocyte infusion, received chemotherapy for a relapse subsequent to allogeneic SCT, and received additional stem cells with or without prior standard dose chemotherapy. The final study cohort was comprised of 110 patients. Detailed information regarding age, sex (donor and recipient), diagnosis, preparative regimen, GVHD prophylaxis, disease status at transplant, cytomegalovirus (CMV) serology, survival status, and causes of death were reviewed for this cohort.

Transplant Regimens

Patients received either myeloablative regimens (cyclophosphamide 120 mg/kg and total body irradiation [TBI] 1200 cGy; cyclophosphamide 120 mg/kg and busulfan 16 mg/kg; cyclophosphamide 7200 mg/m^2, etoposide 2000 mg/m^2, BCNU 400 mg/m^2), reduced intensity (RIC) (fludarabine 90 mg/m^2 and busulfan 8 mg/kg with TBI 400 cGy; fludarabine 90 mg/m^2 and melphalan 140 mg/m^2 with antithymocyte globulin [ATG]) or nonmyeloablative (fludarabine 90 mg/m^2 and TBI 200 cGy) preparative regimens. Stem cell sources were variable, and included peripheral blood stem cells (PBSC), marrow, or cord blood from related or unrelated donors. Standard GVHD prophylaxis regimens with either cyclosporine (CSP) and methotrexate (MTX) or cyclosporine (CSP) and mycophenolate mofetil (MMF) were used. All patients received antimicrobial prophylaxis per institutional guidelines. All patients received prophylaxis for *Pneumocystis carinii* pneumonia after engraftment. Patients were monitored for CMV reactivation either with antigen assays or with polymerase chain reaction (PCR) assays, and were treated in a preemptive manner with ganciclovir or foscarnet.

aGVHD

All patients were graded for aGVHD using the Glucksberg criteria [6] on a weekly basis until day 100, and the maximum grade during this time interval was used. Treatment of aGVHD was variable but the following general principles were used: grade II or higher aGVHD was treated with 2 mg/kg of methylprednisolone or an equivalent dose of prednisone for 7-14 days followed by a taper of 10% every 5-7 days. Patients with isolated upper gut aGVHD could also be treated with 1 mg/kg of methylprednisolone or an equivalent dose of prednisone for 7-14 days followed by a taper of 10% every 5-7 days. Treatment of steroid refractory aGVHD was variable.

cGVHD Diagnosis

Patients with persistent or new diagnosis of GVHD after day 100 were classified as cGVHD and were stratified as limited or extensive [3]. Treatment of cGVHD was variable but followed some general principles: patients with isolated mouth, ocular, or minimal skin cGVHD were treated with topical immunosuppressant (eg, topical steroids). Patients with significant cGVHD were treated with calcineurin inhibitor along with systemic steroids (prednisone). Although duration and dosing of steroids were not standardized, patients typically received treatment until all symptoms of cGVHD were resolved or stabilized and subsequent tapers of immunosuppressive drugs were attempted.

Reclassification of GVHD

A patient was defined as having GVHD after day 100 if a patient had clinical symptoms, signs, or laboratory manifestation of GVHD. Biopsies, when obtained were reviewed. NIH consensus criteria for diagnosis of cGVHD were applied stringently [5]. GVHD after day 100 was reclassified as delayed acute, persistent acute, recurrent acute, classic chronic, overlap chronic. For persistent aGVHD, the assessment was assigned at day 100. For delayed or recurrent aGVHD, the assessment was assigned at onset of manifestation of GVHD. A patient was diagnosed as having cGVHD if there was presence of at least 1 diagnostic clinical sign of cGVHD or presence of at least 1 distinctive manifestation confirmed by pertinent biopsy or other relevant tests and exclusion of other possible diagnoses. If patients had features of aGVHD along with cGVHD, they were classified as overlap cGVHD. Worsening of GVHD was defines as requiring additional immunosuppressive therapy, reescalation of ongoing therapy, or a clinical deterioration attributable to GVHD while on immunosuppressive therapy taper. Data was reviewed serially and time point of peak severity was obtained. At that time point, GVHD was reclassified using the NIH consensus criteria.

Patients who had classic chronic or overlap chronic were scored using the NIH recommended criteria, and a global severity score of mild, moderate, or severe was assigned [5]. Patients with delayed acute, persistent acute, or recurrent acute were not scored, and severity assessments were not assigned. Patients with worsening of GVHD (as defined above) were again reclassified at the time of their peak severity. Severity scores were assigned if they fulfilled criteria for classic cGVHD or overlap cGVHD.

Statistical Analysis

Descriptive statistics, including median and ranges for continuous variables (Table 1) as well as percentages and frequencies for categoric variables were calculated. Groups with nominal outcome were compared using chi-square test; groups with continuous outcomes were compared using Wilcoxon rank sum test. We calculated overall survival (OS) from the day of SCT to the day of death or last follow-up and GVHD-specific survival from the day of onset of GVHD to the day of death or last follow-up. Data were censored for patients alive at their last follow-up visit. Kaplan-Meier survival curves were calculated for each cohort (or clinical group) and were compared using the log rank test.

The old classification stratifies cGVHD as limited and extensive. Recently, the NIH consensus criteria were published. The new criteria stratifies GVHD after day 100 as acute (recurrent, late, or persistent),

Table 1. *Baseline Charactersitics*

	N	%
Patients with GVHD after day 100	73	100
Age, median (years)	42 (range, 1-65)	
Male/female	43/30	59/41
Diagnoses*		
Low risk	40	55
High risk	33	45
Stem cell source		
Marrow	23	32
Peripheral blood stem cells	47	64
Cord blood stem cells	3	4
Preparative regimens		
Myeloablative	56	77
Reduced intensity	6	5
Nonmyeloablative	13	18
HLA match		
HLA identical sibling	60	69
HLA matched unrelated	16	21
Other HLA match	7	10
GVHD prophylaxis		
CSP/MMF	18	25
CSP/MTX	51	70
CSP	4	5

Low risk: CML-CP 1 (n = 6), acute leukemia in CR 1 (n = 20), multiple myeloma (n = 6), low grade NHL (n = 2), MDS (n = 5), CLL (n = 1).

High risk: CML beyond CP1 (n = 1), acute leukemia beyond CR1 (n = 23), NHL other than low grade/Hodgkin disease (n = 8), other = 1.

HLA indicates human leukocyte antigen; CSP, cyclosporine; MTX, methotrexate; MMF, mycophenolate mofetil; GVHD, graft-versus-host disease; NHL, non-Hodgkin lymphoma; CR, complete remission; CLL, chronic lymphocytic leukemia; CP, chronic phase; MDS, myelodysplastic syndromes; CML, chronic myelogenous leukemia.

overlap GVHD (features of both acute and classic) and chronic classic subtypes. Because of small sample size, in some analyses, patients with any acute feature (acute and overlap) were combined and compared with chronic classic. Cox proportional hazard regression analyses using penalized maximum likelihood estimation [7] was used to assess the relationship of old and new GVHD classifications to survival. To perform the most robust analysis of the relationship between GVHD subtype and survival, GVHD subtype was included as a time-dependent variable. The subtype of GVHD was assessed at onset of GVHD (after day 100), and at maximum clinical worsening. The status of GVHD subtype was permitted to vary over transplant to GVHD onset time, onset time to peak time, and peak time to last follow-up. The GVHD subtype was coded as "no GVHD" for the period of SCT to onset time. and coded separately according to the GVHD status of onset and peak time. Statistical assumptions included that a patient with chronic extensive GVHD could not revert back to limited cGVHD. In the new classification, assumptions included that patients with chronic classic could not

revert to only aGVHD. Other covariates included in the Cox model were age at transplant and donor status (related versus unrelated). Cumulative incidence of

NRM was estimated using the Kaplan-Meier method with adjusting for relapse as a competing risk event [8,9]. Analyses were performed using SPSS version 13, SAS system version 9.1 and R version 2.1.1.

RESULTS

The final study cohort consisted of 110 patients. One patient had bronchiolitis obliterans with organizing pneumonia (BOOP) after day 100 without any GVHD, and was unclassifiable. Vanderbilt adult service, associated Veterans Hospital, and the pediatric service contributed 72 (66%), 18 (17%), and 19 (17%) patients, respectively. The median follow-up of the surviving patients was 3.26 years (range: 2.02 to 5.1). The 3-year OS was 50% (95% confidence interval [CI] 0.38 to 0.62). A total of 73 (67%) patients were identified as having GVHD after day 100. Table 1 gives the detailed demographic information, and disease and transplant characteristics of the 73 patients that had GVHD at or after day 100.

No-GVHD Cohort Survival

Thirty-seven (33%) patients did not have any GVHD after day 100. The mean follow-up of patients without GVHD was significantly shorter than patients with GVHD after day 100 (1.8 years versus 2.5 years, $P = .01$). The OS of this group was significantly inferior compared to patients who developed GVHD after day 100 (44% versus 60%, $P = .026$) (Figure 1A). In this analysis, time was measured from day 100 until last follow-up. Twenty of 36 patients (55%) without GVHD died after SCT, 12 (60%) from relapse or progression of their disease, and 8 (40%) from transplant related complications (TRM). Although the causes of death were not significantly different compared with patients with GVHD after day 100, patients without GVHD had a significantly shorter time to death compared to patients with GVHD after day 100 (0.73 years versus 1.27 years, $P = .002$).

GVHD Reclassification

Of the 73 patients with GVHD after day 100, 14 (19%) were classified as limited cGVHD, and 59 (81%) were classified as extensive cGVHD. One pa-
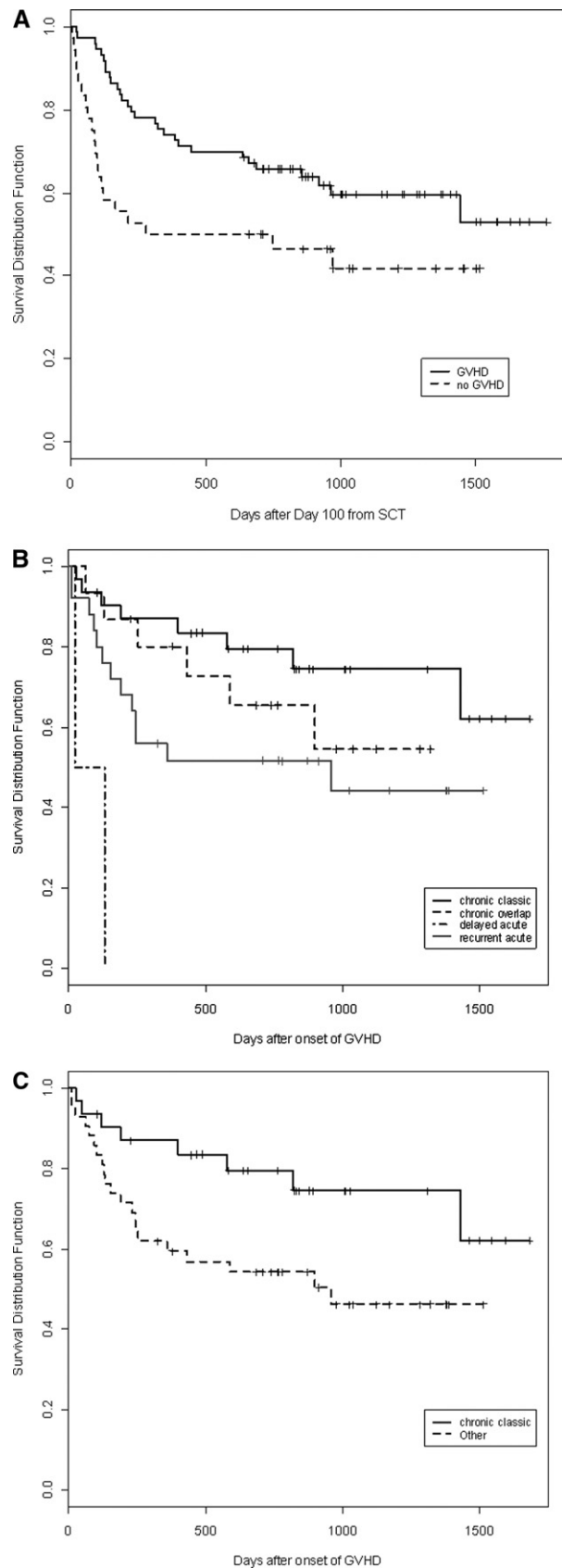
Figure 1. A, OS for the entire cohort: stratified by presence or absence of GVHD after day 100 from SCT. Patients with GVHD after day 100 had a superior survival compared with patients without GVHD (44% versus 60%, $P = .026$). Time is measured from day 100 until last follow-up. B, OS: stratified by GVHD subtype as proposed by NIH consensus criteria. Kaplan-Meier plot shows significant difference in survival of various GVHD subtypes ($P = .0005$). C, OS: classic cGVHD compared to other subtypes (persistent, recurrent, delayed aGVHD, and overlap GVHD). Kaplan-Meier shows superior survival of classic cGVHD compared to other subtypes (3-year OS 68% versus 46%, $P = .033$).

Table 2. *GVHD Reclassification*

	Old Classification Number (%)	
New Classification	**Limited cGVHD** **N = 14**	**Extensive cGVHD** **N = 59**
Persistent aGVHD	1 (7)	2 (3)
Recurrent aGVHD	4 (29)	18 (31)
Delayed aGVHD	—	2 (3)
Overlap GVHD	—	15 (26)
Classic cGVHD	9 (64)	22 (37)

aGVHD indicates acute graft-versus-host disease; cGVHD, chronic graft-versus-host disease.

tient had BOOP without any systemic features of GVHD, and was not classifiable. The 3-year OS of the limited cGVHD cohort was significantly inferior compared to extensive cGVHD (21% versus 63.6%, $P = .037$).

The recently published NIH consensus criteria were used for the 73 patients who were classified as having GVHD after day 100. Based on this classification, GVHD after day 100 was reclassified as persistent aGVHD (3 patients, 4%), recurrent aGVHD (22 patients, 30%), delayed aGVHD (2 patients, 3%), overlap GVHD (features of both acute and classic cGVHD) (15 patients, 20%), and classic cGVHD (31 patients, 42%). One patient with BOOP was not classifiable as any subset of aGVDH or cGVHD. Patients with limited cGVHD (14 patients) were reclassified as persistent acute (1 patient, 7%), recurrent acute (4 patients, 29%), and classic cGVHD (9 patients, 64%) (Table 2). Patients with extensive cGVHD (59 patients) were reclassified as persistent acute (2 patients, 3%), delayed acute (2 patients, 3%), recurrent acute (18 patients, 31%), classic chronic (22 patients, 37%), and overlap GVHD (15 patients, 26%) respectively (Table 2).

Thirty-one (42%) had no worsening and 42 (58%) had worsening of their GVHD, respectively. Among patients with classic cGVHD, 22 of 31 (65%) had worsening of GVHD. This was significantly higher

than other GVHD types where 20 of 42 patients (47%) had worsening of GVHD ($P = .046$).

cGVHD Severity

Of the 15 patients with overlap cGVHD, 5 patients (33%), 9 patients (60%), and 1 patient (7%) were classified as mild, moderate, or severe GVHD. Four (13%), 18 (58%), and 9 (29%) patients with classic cGVHD were classified as mild, moderate, and severe, respectively (Table 3). Using the old classification, clinical worsening was seen in 7 of 14 (50%) patients with limited cGVHD, and 35 of 59 (41%) patients with extensive cGVHD. Nine of 15 patients (60%) with overlap GVHD had worsening compared with 22 of 31 (71%) patients with classic cGVHD ($P = .457$).

GVHD Cohort Survival

OS. Of the 73 patients with GVHD after day 100, 17 patients died from transplant complications and 12 from relapse or progression of the underlying disease for an overall mortality of 40% (Table 4). There was no difference in cause of death (relapse versus NRM causes) in patients with limited compared with cGVHD ($P = .568$). The causes of death were not different when patients with classic cGVHD were compared with other subtypes of GVHD ($P = .487$). Median time to development of GVHD were 89 days, 126 days, 119 days, 159 days, and 244 days for persistent acute, delayed acute, recurrent acute, chronic overlap, and classic cGVHD. The OS for patients with persistent acute, recurrent aGVHD, delayed aGVHD, overlap cGVHD, and classic cGVHD were significantly different ($P = .0005$) (Figure 1B). Three-year OS for patients with persistent acute, recurrent aGVHD, delayed aGVHD, overlap cGVHD, and classic cGVHD were 100%, 45%, 0%, 57%, and 67%, respectively (Figure 1B). This difference in survival (measured from day of SCT) was more apparent when patients with any aGVHD features (recurrent

Table 3. *Distribution of Severity of Overlap and Classic cGVHD at Onset and Peak*

Severity Score at GVHD Onset		Severity Score at GVHD Peak*		Patients with
Overlap GVHD (N = 15)	**Number (%)**	**Overlap GVHD (N = 15)**	**Number (%)**	**Clinical Worsening†**
Mild	5 (33)	**Mild**	1 (7)	9 (60)
Moderate	9 (60)	**Moderate**	9 (60)	
Severe	1 (7)	**Severe**	5 (33)	
Classic cGVHD (N = 31)		**Classic cGVHD (N = 31)**		
Mild	4 (13)	**Mild**	2 (7)	22 (71)
Moderate	18 (58)	**Moderate**	16 (52)	
Severe	9 (29)	**Severe**	13 (41)	

GVHD indicates graft-versus-host disease; cGVHD, chronic graft-versus-host disease.

*Patients with no clinical worsening were assigned severity scores at onset.

†Clinical worsening was defined as increased symptoms or signs (including while on immunosuppressive taper).

Table 4. *Cause of Death among Various GVHD Types*

GVHD Type	N	Cause of Death	
		NRM	RRM
Old Classification			
Limited*	14	4 (29%)	5 (36%)
Extensive*	59	8 (14%)	12 (20%)
NIH Classification			
Persistent aGVHD	3	None	None
Delayed aGVHD	2	1 (50%)	1 (50%)
Recurrent aGVHD	22	6 (27%)	3 (2%)
Overlap GVHD	15	4 (27%)	2 (13%)
Chronic classic GVHD†	31	6 (19%)	2 (6%)

NRM indicates nonrelapse mortality; GVHD, graft-versus-host disease; RRM, death from relapse or progression; aGVHD, acute graft-versus-host disease.

*$P = .568$ comparing cause of death (limited versus extensive), chi-square test.

†$P = .487$ comparing cause of death (classic chronic versus other subtypes), chi-square test.

acute, delayed acute, and overlap chronic) was compared with classic cGVHD (3-year OS 47.2% versus 66.7%, $P = .015$). This effect persisted when survival was measured from onset of GVHD and patients with classic cGVHD (68% versus 46%) had a better GVHD-specific survival (calculated from onset of GVHD after day 100 to last follow-up) compared to other subtypes ($P = .0336$) (Figure 1C).

OSs of patients with mild (9 patients), moderate (27 patients), and severe (10 patients) GVHD at the onset of GVHD diagnosis were not significantly different ($P = .94$) with 3-year OSs of 38.9% versus 73.8% versus 59.3% (Figure 2A). Similarly, the 3-year OS did not differ with respect to the peak severity of the GVHD score. In these patients the 3-year survival was 66.7%, 73.8%, and 59.3% for patients with mild (3 patients), moderate (25 patients), and severe (21 patients) GVHD ($P = .41$) (Figure 2B). Severities were not predictive of survival in patients with either classic cGVHD or overlap cGVHD.

Nonrelapse Mortality (NRM)

The 3-year cumulative NRM with relapse as a competing risk for the entire cohort was 21%. There was no difference in cumulative NRM in patients with GVHD after day 100 compared to patients who had no GVHD after day 100. NRM in patients with chronic classic GVHD was not different compared to patients with other subtypes of GVHD ($P = .57$). Limited or extensive GVHD at onset had no impact on cumulative NRM. Among patients with overlap and chronic classic GVHD, severity at onset or at peak did not influence NRM.

Other Risk Factors

Other known risk factors that predict for survival in patients with cGVHD were analyzed. There was no

difference in survival, either OS or nonrelapse in patients with a platelet count of less than or more than 100×10^9/L at the time of diagnosis of GVHD. Similarly, a bilirubin of less than or more than 3 mg/dL at the time of diagnosis of GVHD had no impact on survival. Source of stem cell had no impact on OS. Patients undergoing matched related donor stem cell transplant had a better survival compared with patients undergoing alternate donor transplants ($P = .029$, 2-year OS: 73.9% versus 47.8%). Age above or below 40 years or as a continuous variable had no impact on survival.

Multivariate Analysis

Cox proportional multivariate models were constructed using the old and the revised classification.
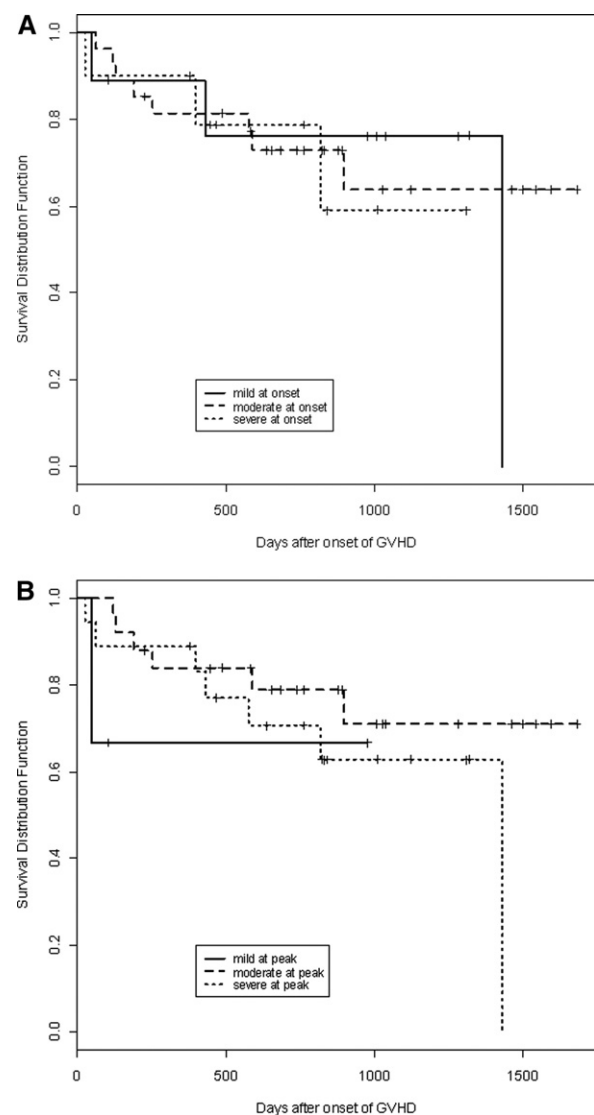
Figure 2. A, OS: stratified by severity at onset of GVHD (classic and overlap only) after day 100. Kaplan-Meier shows no difference in survival ($P = .94$). B, OS: stratified by peak severity of GVHD (classic and overlap only) after day 100. Kaplan-Meier shows no difference in survival ($P = .94$).

Table 5. *Multivariable Analyses*

Variable	Hazard Ratio (95% CI)	P-value
Old Classification		
Limited cGVHD	1	
Extensive cGVHD	0.35 (0.16, 0.80)	.015
Unrelated donor	1	
Related donor	0.42 (0.19, 0.99)	.049
Age at transplant	1.03 (0.97, 1.02)	.813
Old Classification (Accounting for Subtype Changes)		
Limited cGVHD	1	
Extensive cGVHD	0.36 (0.16, 0.91)	.033
Unrelated donor	1	
Related donor	0.46 (0.20, 1.07)	.071
Age at transplant	1.002 (0.97, 1.02)	.89
NIH Consensus Classification		
Classic cGVHD	1	
Acute feature of GVHD*	1.72 (0.71, 4.40)	.228
Unrelated donor	1	
Related donor	0.61 (0.26, 1.49)	.279
Age at transplant	1.00 (0.98, 1.02)	.781
NIH Consensus Classification (Accounting for Subtype Changes)		
Classic cGVHD	1	
Acute feature of GVHD*	3.36 (1.25, 11.09)	.0144
Unrelated donor	1	
Related donor	0.66 (0.29, 1.54)	.333
Age at transplant	1.007 (0.98, 1.03)	.555

GVHD indicates graft-versus-host disease; cGVHD, chronic graft-versus-host-disease.

*Includes any acute feature of GVHD after day 100 (late acute, recurrent acute, persistent acute, overlap GVHD).

Age at transplant and type of donor were adjusted for in all models (Table 5). GVHD was analyzed as a time-dependent variable. When adjusted for donor status and age, extensive GVHD was associated with improved survival with a hazard ratio (HR) of 0.35 (95% CI 0.16 to 0.80, $P = .015$). Adjusted for donor status and age at transplant, survival associated with any acute feature of GVHD after day 100 (includes late acute, delayed acute, recurrent acute, overlap) was not statistically different compared with classic cGVHD (HR 1.72, 95% CI 0.71 to 4.40, $P = .2283$).

Clinically, patients can transition from one subtype of GVHD to another. To address these issues multivariate models were constructed using rules. Patients were allowed to transition from one subtype to another, except patients with extensive cGVHD could not transition back to limited cGVHD, and classic or overlap GVHD could not transition to acute-only GVHD. Overlap GVHD and acute-only GVHD were combined as 1 group because of sample size limitations. Adjusted for donor status and age at transplant, extensive cGVHD (accounting for subtype changes) was associated with an HR of 0.36 (95% CI 0.16 to 0.91, $P = .033$). Any acute feature of GVHD (accounting for subtype changes) after day 100 (includes late acute, recurrent acute, persistent acute, overlap GVHD) was

associated with an HR of 3.36 (95% CI 1.25 to 11.09, $P = .0144$) when compared to classic cGVHD.

DISCUSSION

In this study we show the importance and the impact of the recently published NIH consensus criteria on classification of cGVHD. Patients with classic cGVHD have the best survival compared to patients with other subtypes of GVHD after day 100. Although, severity of GVHD may be important in decision making regarding systemic immunosuppression, we could not show that it had an impact on survival.

GVHD after day 100 from transplant has historically been classified as cGVHD. In the past several years, with advances in nonmyeloablative transplant and RIC regimens, the natural history of aGVHD has changed, with a higher proportion of patients being diagnosed with late-onset (after day 100) aGVHD [10,11]. Also, manifestations of aGVHD and cGVHD are known to exist simultaneously. This makes the arbitrary definition of any GVHD after day 100 being called cGVHD invalid. The most common classification system in use stratifies cGVHD into limited and extensive [3]. This classification was based on a report of 20 patients. The prognostic significance of this classification is not known. In our study, patients with limited (14 patients) and extensive (59 patients) cGVHD had a 3-year OS of 21% and 63%, respectively ($P = .037$). This classification is not reproducible or prognostic for late TRM. The incidence of cGVHD is variable among studies, and is influenced by a variety of nonmodifiable (age of patient, age of donor, HLA match) or modifiable factors (GVHD prophylaxis, source of stem cells, graft manipulation) [12-16]. Aside from these, interobserver variability, lack of standardized diagnostic criteria, and the protean manifestations of cGVHD have impaired the development of a common platform for diagnosis and classification of cGVHD. Recently, the NIH consensus development project on criteria for clinical trials in cGVHD published the diagnosis and staging working group report [5]. The incidence and outcome of the various subtypes of GVHD and the impact of the newly developed scale for global assessment of severity of cGVHD is unknown. We studied a cohort of patients transplanted in a 2-year time interval at a single center. Although limited by its retrospective nature and recording bias, the patient cohort was homogenous as patients receiving second transplant, chemotherapy after transplant, or donor lymphocyte infusion after transplant were excluded. Because the natural history of aGVHD prior to day 100 is well known, our cohort included patients who survived at least until day 100. Survival of the patients signifi-

cantly differed depending on subtype of GVHD after day 100. Patients with classic cGVHD had the best outcome with a 68% 3-year OS. Patients with other subtypes of GVHD had a poor survival, with the worst outcome in patients with recurrent aGVHD and delayed aGVHD. Thus, our study identifies a high-risk group of patients among patients who present with GVHD after day 100 from SCT. Patients with extensive cGVHD had significantly less risk of death compared to limited cGVHD (HR 0.284; 95% CI 0.12-0.67). The causes of death (relapse versus NRM) were not significantly different between the 2 groups ($P =$ .568, Table 4). This is contradictory compared to previously published reports [17]. Possible explanations for this contradictory result could include small number of limited cGVHD (n = 14) to study the impact of this form of GVHD on survival. Another possible explanation is that in the previously published study [17] the rate of misclassification of extensive cGVHD as limited cGVHD study was reported as 65% to 67% for sibling transplants and 43% (83% of all cases) for unrelated transplants. The authors did not reclassify these patients because database limitations prevented the application of limited and extensive criteria with certainty. In our study, all interval data was available for stringent application of limited and extensive cGVHD criteria, and thus misclassification was minimized. Although OS of patients with classic cGVHD was significantly better than other GVHD subtypes, the specific reason for this survival benefit could not be determined because of small sample size.

GVHD subtypes were assigned at onset and at clinical worsening. Clinically, it is common for the GVHD subtype to change over follow-up. Multivariable analyses accounting for these changes in subtype, continued to show that extensive cGVHD was associated with a superior survival (HR 0.36, $P =$ .033) when adjusted for donor status and age at transplant. Similarly, accounting for GVHD subtype variation during clinical course, any acute feature of GVHD after day 100 (includes late acute, recurrent acute, persistent acute, overlap GVHD) was associated with an inferior survival (HR 3.36, $P =$ .0144) compared with classic cGVHD.

The NIH consensus report has suggested applying global assessment of severity of GVHD after a diagnosis of cGVHD has been made [5]. We applied this scale only to patients with classic cGVHD and overlap cGVHD, as most of the individual organ function scoring on which the scale is based is targeted toward cGVHD symptoms. In our cohort of patients, severity at onset or at peak of cGVHD activity did not correlate with outcome. The global severity score as proposed by the NIH consensus criteria may be useful as a descriptive clinical indicator of morbidity and the need for immunosuppressive therapy, but within the limits of our study, we did not find that this score has prognostic value with respect to survival.

Patients with any features of aGVHD after day 100 had the poorest OS compared with classic cGVHD (3-year OS 47.2% versus 66.7%, $P =$ .015). This patient population needs to be preferentially enrolled on clinic trials of GVHD, exploring novel agents. Patients with classic cGVHD have a 3-year OS of 67%, and could preferentially be enrolled in studies targeting topical treatment and organ-specific approaches. These patients should continue to be followed for late secondary complications of stem cell transplant. The NIH consensus criterion represents a true advance in subclassifying GVHD after day 100, and provides a common platform for clinical trials. Our study validated the importance of subclassifying GVHD after day 100 as proposed by the NIH consensus criterion. It is important to validate the proposed NIH criteria in a prospective fashion in a multicenter setting.

REFERENCES

1. Sullivan K. Graft vs. host disease. In: Blume KG, Forman SJ, Appelbaum F, eds. *Thomas' Hematopoietic Cell Transplantation.* Oxford, UK: Blackwell Publishing, 2006:635-664.
2. Sullivan KM, Witherspoon RP, Storb R, et al. Alternating-day cyclosporine and prednisone for treatment of high-risk chronic graft-v-host disease. *Blood.* 1988;72:555-561.
3. Shulman HM, Sullivan KM, Weiden PL, et al. Chronic graft-versus-host syndrome in man. A long-term clinicopathologic study of 20 Seattle patients. *Am J Med.* 1980;69:204-217.
4. Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2003;9:215-233.
5. 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;11:945-956.
6. Glucksberg H, Storb R, Fefer A, et al. Clinical manifestations of graft-versus-host disease in human recipients of marrow from HL-A-matched sibling donors. *Transplantation.* 1974;18:295-304.
7. Heinze G, Schemper M. A solution to the problem of monotone likelihood in Cox regression. *Biometrics.* 2001;57:114-119.
8. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat.* 1988;16:1141-1154.
9. 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;18:695-706.

10. Mielcarek M, Martin PJ, Leisenring W, et al. Graft-versus-host disease after nonmyeloablative versus conventional hematopoietic stem cell transplantation. *Blood*. 2003;102:756-762.

11. Flowers ME, Parker PM, Johnston LJ, et al. Comparison of chronic graft-versus-host disease after transplantation of peripheral blood stem cells versus bone marrow in allogeneic recipients: long-term follow-up of a randomized trial. *Blood*. 2002;100:415-419.

12. Storb R, Prentice RL, Sullivan KM, et al. Predictive factors in chronic graft-versus-host disease in patients with aplastic anemia treated by marrow transplantation from HLA-identical siblings. *Ann Intern Med*. 1983;98:461-466.

13. Niederwieser D, Pepe M, Storb R, et al. Factors predicting chronic graft-versus-host disease and survival after marrow transplantation for aplastic anemia. *Bone Marrow Transplant*. 1989;4:151-156.

14. Atkinson K, Horowitz MM, Gale RP, et al. Risk factors for chronic graft-versus-host disease after HLA-identical sibling bone marrow transplantation. *Blood*. 1990;75:2459-2464.

15. Ochs LA, Miller WJ, Filipovich AH, et al. Predictive factors for chronic graft-versus-host disease after histocompatible sibling donor bone marrow transplantation. *Bone Marrow Transplant*. 1994;13:455-460.

16. Carlens S, Ringden O, Remberger M, et al. Risk factors for chronic graft-versus-host disease after bone marrow transplantation: a retrospective single centre analysis. *Bone Marrow Transplant*. 1998;22:755-761.

17. Lee SJ, Klein JP, Barrett AJ, et al. Severity of chronic graft-versus-host disease: association with treatment-related mortality and relapse. *Blood*. 2002;100:406-414.