Association of severity of organ involvement with mortality and recurrent malignancy in patients with chronic graft-*versus*-host disease

Yoshihiro Inamoto,[1] Paul J. Martin,[1] Barry E. Storer,[1] Jeanne Palmer,[2] Daniel J. Weisdorf,[3] Joseph Pidala,[4] Mary E. D. Flowers,[1] Mukta Arora,[3] Madan Jagasia,[5] Sally Arai,[6] Xiaoyu Chai,[1] Steven Z. Pavletic,[7] Georgia B. Vogelsang,[8] and Stephanie J. Lee[1] on behalf of the Chronic GVHD Consortium

[1]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA; [2]Division of Hematology and Oncology, Mayo Clinic Arizona, Scottsdale, AZ, [3]Blood and Marrow Transplant Program, University of Minnesota, Minneapolis, MN; [4]Blood and Marrow Transplantation, Moffitt Cancer Center, Tampa, FL; [5]Hematology and Stem Cell Transplant Program, Vanderbilt University Medical Center, Nashville, TN; [6]Division of Blood and Marrow Transplantation, Stanford University Medical Center, Stanford, CA; [7]National Cancer Institute, Bethesda, MD; and [8]Oncology, Johns Hopkins Hospital, Baltimore, MD, USA

ABSTRACT

The National Institutes of Health global score for chronic graft-*versus*-host disease was devised by experts but was not based on empirical data. We hypothesized that analysis of prospectively collected data would enable derivation of a more accurate model for estimating mortality risk. We analyzed 574 adult patients with chronic graft-*versus*-host disease enrolled in a multicenter, observational study, using multivariate time-varying analysis accounting for serial changes in severity of involvement of eight individual organ sites over time. In the training set, severity of skin, mouth, gastrointestinal tract, liver and lung involvement were independently associated with the risk of non-relapse mortality. Weighted mortality points were assigned to individual organs based on the hazard ratios and were summed. The population was divided into three risk groups based on the total mortality points. The three new risk groups were validated in an independent validation set, but did not show better discriminative performance than the National Institutes of Health global score. As compared to a moderate or mild global score, a severe global score was associated with increased risks of non-relapse and overall mortality across time but not with a decreased risk of recurrent malignancy. The National Institutes of Health global score predicts patients' mortality risk throughout the course of their chronic graft-*versus*-host disease. Further research is required in order to improve outcomes in patients with severe chronic graft-*versus*-host disease, since their risk of mortality remains elevated.

Introduction

Chronic graft-*versus*-host disease (GVHD) occurs in 30-50% of patients after allogeneic hematopoietic cell transplantation and is a leading cause of late morbidity and mortality,[1,2] although it is also associated with a decreased risk of recurrent malignancy.[3-7] The 2005 National Institutes of Health (NIH) consensus criteria proposed a scoring algorithm for overall severity (NIH global score) of GVHD based on the severity of manifestations in eight individual organs (NIH organ severity score).[8] Scores may be assessed at any point after diagnosis of chronic GVHD because manifestations of chronic GVHD vary across time. Previous studies using the older definition of chronic GVHD based on time after transplantation showed that the overall severity of chronic GVHD at diagnosis was associated with risk of subsequent non-relapse mortality but was not associated with risk of recurrent malignancy.[9] Although the NIH global score was developed through expert opinion, several studies have shown that the global score at onset of chronic GVHD is associated with risk of subsequent mortality.[10-13]

Many issues do, however, remain to be determined. (i) Since the NIH global score was based on expert opinion and was not

originally intended to predict mortality, does this score provide an optimal model for predicting mortality risk in patients with chronic GVHD? We hypothesized that empirically derived estimates of overall mortality risk incorporating the relative importance of different organ involvement might be more accurate than estimates derived from the opinion-based global score. (ii) Does the NIH global score predict mortality risk when it is applied at time points after the onset of treatment for chronic GVHD? (iii) Does the NIH global score correlate with risk of recurrent malignancy? To address these important clinical questions, we performed a detailed analysis of data collected in a prospective, multicenter, longitudinal, observational study of patients with chronic GVHD.

Methods

Study cohort

Patients receiving systemic treatment for chronic GVHD were eligible for a prospective observational study by the Chronic GVHD Consortium.[14] The chronic GVHD was diagnosed and scored according to NIH consensus criteria.[8] At enrollment and every 6 months thereafter, clinicians and patients reported standardized information

regarding organ manifestations of chronic GVHD. Patients enrolled ≤3 months after the diagnosis of chronic GVHD had an additional assessment at 3 months after enrollment. Patients were treated according to institutional practice. A total of 2271 visits from 574 adult patients through to January 2013 were included in the analysis. Thirty-three visits with missing organ severity scores in at least one organ were excluded from the analyses. Four hundred and twenty-four patients with 1602 visit ratings (71%) were randomly selected as a training set in order to develop a new mortality model, and the remaining 150 patients with 600 visit ratings were reserved as a validation set. The study was approved by the Institutional Review Board of each participating center, and participants gave written informed consent in accordance with the Declaration of Helsinki.

Statistical analysis

Longitudinal scores of GVHD severity for individual organs were correlated with the risks of non-relapse mortality, overall mortality and recurrent malignancy. In this "dynamic" model, severity scores for individual organs changed over time, and patients were allowed to move to the respective severity group at each research assessment. Time-varying Cox models were used to account for serial changes in disease severity over time, and left truncation was used to account for varied times from initial systemic treatment to entry into the study. Mortality models were adjusted by time after transplantation and known chronic GVHD mortality risk factors not related to specific organs.[15-17] Models for recurrent malignancy were adjusted by time after transplantation and disease risk. Time after transplantation was included in all models instead of incident *versus* prevalent status. Transplant center was also included in all models in order to adjust for potential differences in clinical management.

A backward stepwise selection was used to retain organs that contained at least one statistically significant category at *P* values <0.05. To create an easily applicable mortality model, β coefficients from the final model were multiplied by 1.5 and rounded to yield the lowest integer values. The sum of these integer weights was used to yield the total mortality points. The Akaike information criterion (AIC) was used to compare predictive capability between the models.[18] When two predictive models using the same population have the same number of groups, a lower AIC indicates a better model fit and better predictive capability. Although the models have the same degrees of freedom, a reduction of AIC by 3.8 corresponds to a one degree of freedom likelihood ratio test at the 0.05 significance level and suggests a statistically better model. Smoothed event hazard rates were plotted to illustrate the dynamic change in event rates across time after initial systemic treatment, using the kernel-density estimation.[19]

Results

Patients' characteristics

The median age of the patients was 52 years (range, 19-79 years). Five hundred and six patients (88%) received mobilized blood cell grafts, and 329 (57%) were prepared with myeloablative conditioning regimens. The median follow-up time of survivors was 36.9 months (range, 4.1-82.3 months) after the initial systemic treatment of chronic GVHD. The median time between visits was 5.7 months (range, 1.2-24 months), and 1391 of 1697 follow-up visits (82%) were made at 6 months after the prior visit. Other demographics are summarized in Table 1. The characteristics of the patients in the training and validation sets were similar.

Table 1. Patients' characteristics.

Characteristic	Training set	Validation set	P
N. of patients	424	150	
Median patients' age, range, years	52 (19-79)	54 (19-74)	0.27
Patients' gender, n. (%)			0.33
Male	251 (59)	82 (55)	
Female	173 (41)	68 (45)	
Patients' race, n. (%)			0.12
White	377 (89)	140 (93)	
Others	47 (11)	10 (7)	
Disease, n. (%)			0.18
Acute myeloid leukemia	132 (31)	60 (40)	
Myelodysplastic syndrome	69 (16)	20 (13)	
Chronic myeloid leukemia	18 (4)	12 (8)	
Acute lymphoblastic leukemia	49 (12)	13 (9)	
Malignant lymphoma	81 (19)	21 (14)	
Chronic lymphocytic leukemia	34 (8)	12 (8)	
Multiple myeloma	20 (5)	9 (6)	
Severe aplastic anemia	5 (1)	1 (<2)	
Other	16 (4)	2 (1)	
Disease stage*, n. (%)			0.51
Early	131 (31)	53 (35)	
Intermediate	185 (44)	62 (41)	
Advanced	104 (24)	35 (23)	
Not available	4 (<1)	0 (0)	
Center, n. (%)			0.26
Fred Hutchinson Cancer Research Center	195 (46)	58 (39)	
Stanford University Medical Center	49 (12)	25 (17)	
Dana-Farber Cancer Institute	44 (10)	21 (14)	
University of Minnesota	43 (10)	18 (12)	
Vanderbilt University Medical Center	39 (9)	9 (6)	
Lee Moffitt Cancer Center	26 (6)	12 (8)	
Medical College of Wisconsin	17 (4)	6 (4)	
Memorial Sloan-Kettering Cancer Center	8 (2)	0 (<1)	
Washington University in St. Louis	3 (<1)	1 (<1)	
Graft source, n. (%)			0.43
Bone marrow	33 (8)	7 (5)	
Mobilized blood cells	371 (88)	135 (90)	
Cord blood	20 (5)	8 (5)	
Donor and HLA matching, n. (%)			0.38
HLA-matched related	173 (41)	65 (43)	
HLA-matched unrelated	177 (42)	66 (44)	
HLA-mismatched related	9 (2)	4 (3)	
HLA-mismatched unrelated	60 (14)	15 (10)	
Not available	5 (1)	0 (0)	
Conditioning regimen, n. (%)			0.49
Myeloablative	241 (57)	88 (59)	
Non-myeloablative / reduced-intensity	181 (43)	60 (40)	
Not available	2 (<1)	2 (1)	
Donor-patient gender combination, n. (%)			0.58
Female to male	128 (30)	39 (26)	
Other	292 (69)	110 (73)	
Not available	4 (<1)	1 (<1)	
Donor-patient CMV status, n. (%)			0.74
Either positive	278 (66)	99 (66)	
Both negative	141 (33)	48 (32)	
Not available	5 (1)	3 (2)	
Prior grades II-IV acute GVHD, n. (%)	287 (68)	91 (61)	0.12

*HLA: human leukocyte antigen; CMV: cytomegalovirus; *Early disease included acute leukemia in first remission, chronic myeloid leukemia in chronic phase, myelodysplastic syndrome without excess blasts and non-malignant diseases; advanced disease included acute leukemia in third or later remission, malignant disease not in remission and chronic myeloid leukemia in blast crisis; all other diseases and stages were categorized as intermediate risk.*

Associations of organ severity scores with risk of non-relapse mortality

In the training set, multivariate models showed that organ severity scores in the skin, mouth, gastrointestinal tract, liver and lung were independently associated with the risk of non-relapse mortality (Table 2), while scores for the eyes, joints or fascia and genital tract were not. Weighted mortality points were assigned to individual organs according to the observed β coefficients from the final model. Three risk groups were defined based on the sum of the mortality points (Table 3 left). Patients in both the intermediate-risk category [hazard ratio (HR) 3.56; 95% confidence interval (95% CI), 1.41-8.96; P=0.007] and the high-risk category (HR 16.4; 95% CI, 6.52-41.3; P<0.001) had a higher risk of non-relapse mortality compared with those in the low-risk category. Considering the NIH global score instead, patients in the severe GVHD category, but not those with moderate disease, had a higher risk of non-relapse mortality than patients with mild

GVHD (HR 11.3; 95% CI, 2.67-47.6; P=0.001). As expected, the new mortality model showed better predictive capability compared to the global score model (AIC 532.0 *versus* 545.7, corresponding to a difference of at least three degrees of freedom), since the new model was developed in this training set.

The same scoring algorithm was applied in the independent validation set of 150 patients. Different assignments of risk categories between the new mortality model and the NIH global score were observed in 219 of 622 visits (35%) (*Online Supplementary Table S1*). Risk categories assigned by the NIH global score were often higher than those assigned by the new mortality model. Differences were observed most frequently in 123 visits (20%) classified as moderate risk by the global score but as low risk by the new mortality model (*Online Supplementary Table S2*). Score 2 in the skin (n=61) and liver (n=22) and score 1 in more than two organs (n=21) were the major reasons accounting for moderate global score. The second most

Table 2. Associations of organ scores with risk of non-relapse mortality in the training set.

Organ score	N. of patients*	Univariate HR[†] (95% CI)	P	Multivariate HR[†] (95% CI)	P	Mortality points[‡]
Skin						
0	287	1.00 (reference)		1.00 (reference)		0
1	190	1.51 (0.71-3.22)	0.28	1.54 (0.71-3.35)	0.28	0
2	183	1.78 (0.81-3.91)	0.15	1.96 (0.87-4.39)	0.10	0
3	110	6.25 (3.01-13.0)	<0.001	5.19 (2.40-11.2)	<0.001	2
Mouth						
0	298	1.00 (reference)		1.00 (reference)		0
1	299	2.01 (1.08-3.72)	0.027	1.62 (0.86-3.08)	0.14	0
2-3[§]	89	4.53 (1.93-10.6)	0.001	2.74 (1.05-7.16)	0.04	1
Eye						
0	273	1.00 (reference)				
1	249	0.83 (0.43-1.58)	0.57			
2	147	0.92 (0.44-1.94)	0.84			
3	34	1.70 (0.49-5.97)	0.41			
Gastrointestinal tract						
0	389	1.00 (reference)		1.00 (reference)		0
1	174	2.40 (1.25-4.60)	0.008	2.31 (1.16-4.60)	0.017	1
2-3[§]	50	2.23 (0.74-6.73)	0.16	1.17 (0.36-3.83)	0.79	1
Liver						
0	331	1.00 (reference)		1.00 (reference)		0
1	215	1.14 (0.58-2.24)	0.71	1.12 (0.56-2.25)	0.74	0
2	99	1.32 (0.53-3.24)	0.55	0.94 (0.35-2.49)	0.90	0
3	39	5.10 (1.89-13.8)	0.001	4.44 (1.55-12.8)	0.006	2
Lung						
0	304	1.00 (reference)		1.00 (reference)		0
1	215	2.63 (1.26-5.49)	0.01	2.41 (1.11-5.26)	0.027	1
2	99	4.15 (1.97-8.76)	<0.001	4.58 (2.13-9.86)	<0.001	2
3	39	9.11 (2.79-29.7)	<0.001	8.20 (2.41-27.9)	0.001	3
Joint/fascia						
0	356	1.00 (reference)				
1	187	0.75 (0.34-1.63)	0.46			
2-3[§]	81	1.16 (0.43-3.09)	0.77			
Genital tract[¶]						
0	149	1.00 (reference)				
1	60	0.73 (0.16-3.40)	0.69			
2-3[§]	41	1.59 (0.40-6.28)	0.51			

*Total number of patients contributing to the score at one or more visits. [†]Models were adjusted for time after transplantation, transplant center, patients' age, stem cell source, disease risk, cytomegalovirus status, HLA and donor type, gender mismatch, conditioning intensity, prior acute GVHD and thrombocytopenia at the visit. [‡]β coefficients from the model were multiplied by 1.5 and rounded to the nearest integer. [§]Scores 2 and 3 were combined since score 3 was documented in small numbers of patients. [¶]Female patients.

frequent differences were observed in 73 visits (12%) classified as severe risk by the global score but as intermediate risk by the new mortality model. A lung score of 2 (n=27) and a skin score of 3 (n=22) were the major reasons accounting for a global score denoting severe GVHD.

In the validation set (Table 3 right), both patients in the intermediate-risk category (HR 4.56; 95% CI, 1.12-18.6; P=0.03) and those in the high-risk category (HR 10.1; 95% CI, 2.18-46.6; P=0.003) had higher risks of non-relapse mortality compared with those in the low-risk category according to the new mortality model. Using the NIH global score instead, patients in the severe category had a higher risk of non-relapse mortality compared with those in the mild category (HR 15.3; 95% CI, 1.81-129; P=0.012), but those in the moderate category did not. Based on the AIC, however, the new mortality model showed a significantly worse prediction capability than did the global score model (AIC 202.1 *versus* 193.8, corresponding to a difference of approximately two degrees of freedom).

Associations of organ severity scores with risk of overall mortality

Associations of organ severity scores with risk of overall mortality were examined in all 574 patients (*Online Supplementary Table S3*). Multivariate models showed that organ severity scores for the skin, gastrointestinal tract, liver and lung were independently associated with overall mortality, consistent with the analysis of non-relapse mortality. Organ severity scores for the mouth, eyes, joints or fascia and genital tract did not show statistically significant associations with risk of overall mortality.

Associations of organ severity scores with risk of recurrent malignancy

Associations of organ severity scores with risk of recurrent malignancy were examined in 565 patients who had undergone transplantation for malignant diseases (*Online Supplementary Table S4*). An eye score of 2 or 3 was associated with a decreased risk of recurrent malignancy compared with an eye score of 0 (HR 0.38; 95% CI, 0.18-0.77; P=0.007), but the severity scores for other organs did not show statistically significant associations with risk of recurrent malignancy. We found no statistical interaction of disease type or disease risk in these analyses. Results were similar even when involvement *versus* no involvement was considered for analysis (*Online Supplementary Table S5*).

Change in hazard of mortality and recurrent malignancy over time according to the National Institutes of Health global score

In order to investigate whether the risk associated with a severity group changed over time, smoothed hazard rates were plotted for the entire cohort according to NIH global score (Figure 1). In this dynamic analysis, risk categories were revised at each change in chronic GVHD severity. Rates of non-relapse mortality decreased gradually over time after initial treatment, but remained clearly distinguishable by the global score (Figure 1A). For example, among patients classified as having a severe global score at 1 year after initial treatment, the estimated non-relapse mortality rate per patient-year is approximately 18%. The corresponding estimate at 4 years after initial treatment is approximately 7% per patient-year. In contrast, the non-relapse mortality rate is below 5% per patient-year across time for patients with a mild or moderate global score. The curve was limited for patients with a mild global score due to the small number of events. Overall mortality rates were also stratified clearly across time by the NIH global score (Figure 1B). In contrast, recurrent malignancy rates were similar across time regardless of global score after initial treatment (Figure 1C). These observations were also confirmed by Cox models (Table 4). We found no statistical interaction of disease type or disease risk in the analyses of recurrent malignancy.

Discussion

In our analysis of 562 patients with chronic GVHD assessed longitudinally, the association of organ severity with risk of mortality differed according to individual organ sites. The strongest associations were observed for the lung, followed by the liver, skin, gastrointestinal tract and mouth. We attempted to develop a better mortality model based on these weighted associations, but the validation set analysis showed that the new model performed less well than the NIH global score. Given the simplicity of the NIH global score and its increasing use in clinical research and practice since the 2005 NIH Consensus Conference, we conclude that the NIH global score serves as an adequate model for predicting mortality risk throughout the course of chronic GVHD in patients with this condition.

While the NIH global score was originally intended to

Table 3. Associations of risk groups with risk of non-relapse mortality.

Model	Total mortality points	N. of patients*	Training set HR† (95% CI)	P	N. of patients*	Validation set HR† (95% CI)	P
New mortality model							
Low	0	251	1.00 (reference)		98	1.00 (reference)	
Intermediate	1-2	328	3.56 (1.41-8.96)	0.007	119	4.56 (1.12-18.6)	0.034
High	≥3	156	16.4 (6.52-41.3)	<0.001	47	10.1 (2.18-46.6)	0.003
NIH global score							
Mild	NA	140	1.00 (reference)		56	1.00 (reference)	
Moderate	NA	334	2.67 (0.60-11.9)	0.20	121	4.29 (0.47-38.8)	0.20
Severe	NA	241	11.3 (2.67-47.6)	0.001	79	15.3 (1.81-129)	0.012

HR: hazard ratio; CI: confidence interval; NA: not applicable. *Total number of patients contributing to the category at one or more visits. †Models were adjusted for time after transplantation, transplant center, patients' age, stem cell source, disease risk, cytomegalovirus status, HLA and donor type, gender mismatch, conditioning intensity, prior acute GVHD and thrombocytopenia at the visit.
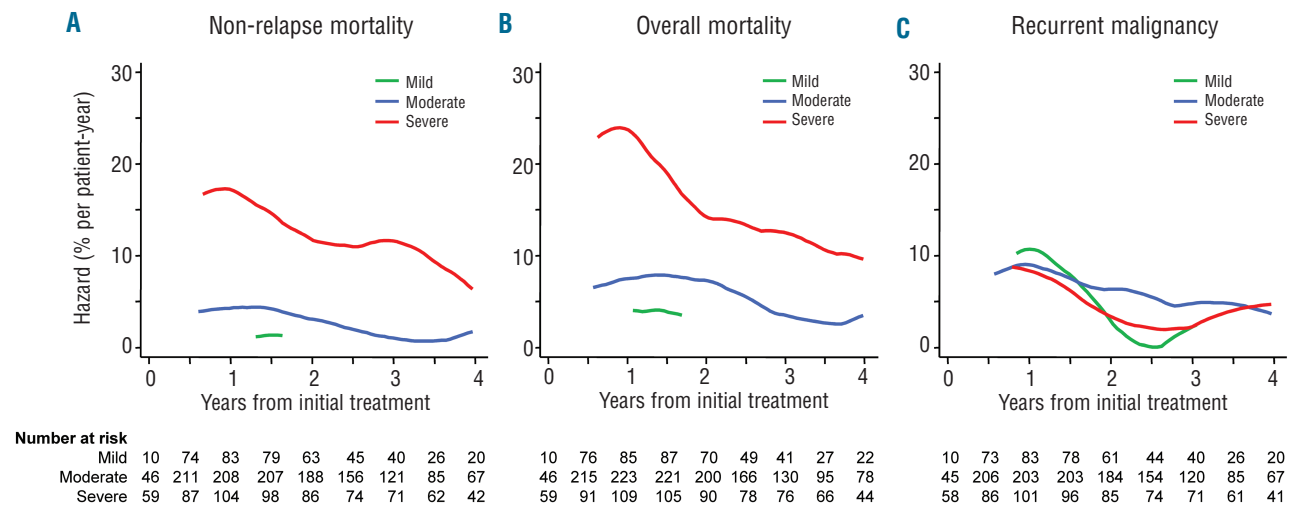
Figure 1. Smoothed hazard rates over time according to NIH global score. Plots are not illustrated for the first 6 months or for time intervals with small numbers of events. (**A**) Non-relapse mortality, (**B**) overall mortality, (**C**) recurrent malignancy.

Table 4. Associations of NIH global score with outcomes.

NIH global score	Non-relapse mortality			Overall mortality			Recurrent malignancy		
	N. of patients*	HR[†] (95% CI)	P	N. of patients*	HR[†] (95% CI)	P	N. of patients*	HR[‡] (95% CI)	P
Mild	196	1.00 (reference)		196	1.00 (reference)		194	1.00 (reference)	
Moderate	455	4.27 (0.99-18.4)	0.051	455	2.79 (1.24-6.30)	0.013	447	1.26 (0.68-2.35)	0.46
Severe	320	17.1 (4.12-71.3)	<0.001	320	7.59 (3.42-16.9)	<0.001	314	1.27 (0.64-2.52)	0.49

*Total number of patients contributing to the category at one or more visits. †Models were adjusted for time after transplantation, transplant center, patients' age, stem cell source, disease risk, cytomegalovirus status, HLA and donor type, gender mismatch, conditioning intensity, prior acute GVHD and thrombocytopenia at the visit. ‡Models were adjusted for time after transplantation, transplant center and disease risk.

predict functional disability due to chronic GVHD and need for treatment, previous studies showed that global score of chronic GVHD at onset was associated with subsequent mortality.[10-13] We extended these findings by taking the dynamic changes in chronic GVHD severity into account, which mirrors the clinical reality of episodic improvement and worsening over time. Hazard analysis confirmed that the NIH global score correlated well with mortality rates regardless of time since initial treatment, suggesting that the NIH global score is applicable to patients throughout the course of the disease after initial treatment. In general, mortality rates decreased over time after initial treatment, but patients with a severe global score continued to have substantially increased rates of non-relapse mortality compared to those with mild or moderate global scores. The poor prognosis of patients with severe chronic GVHD emphasizes the need for more research to improve outcomes in this risk group.

The current study elucidated independent associations of severity of organ involvement with mortality, while previous studies examined associations of single organ involvement with mortality.[9,20-24] One study using the older definition of chronic GVHD showed that oral involvement at onset was associated with a decreased risk of subsequent mortality,[9] but our results did not confirm this observation. It is difficult to explain the biological mechanisms for the association of an eye score of 2 or 3 with a decreased risk of recurrent malignancy. One study found better survival rates

among patients with ocular GVHD than those without ocular GVHD, but recurrent malignancy was not examined in that study.[25] Additional studies are warranted to determine whether this observation holds true in other cohorts.

This study has some limitations. First, the analytic power might not have been sufficient to observe mortality associated with severe involvement in some organ sites. Since such severe involvement was rare in our contemporary cohort, further accumulation of severe cases is required to address these subpopulations adequately. Second, the analysis was structured according to clinic visits. Although 82% of follow-up visits were made at 6 months after the prior visit, the analysis might not account for potentially important changes in chronic GVHD severity between successive visits. Third, most patients received mobilized blood cell grafts and results may not apply to patients who receive bone marrow or cord blood grafts. Lastly, as recommended by the NIH criteria,[8] the current organ severity scores describe a patient's current condition without considering the attribution of the manifestations and without distinguishing active disease from fixed deficits. Future studies could examine mortality algorithms that incorporate these additional considerations.

The novelty of this study comes from incorporating the waxing and waning manifestations of chronic GVHD from eight organ sites into a single model. By using time-varying Cox models accounting for changes in severity of organ involvement over time, we were able to analyze a

complex clinical syndrome and examine the association of organ scores with outcome risks in a way that is applicable throughout the course of chronic GVHD. In addition, hazard analysis illustrated dynamic changes in event rates over time according to serial assessments of overall severity of chronic GVHD. The results support the utility of the NIH global score to predict mortality risk at any time after beginning treatment, even if the severity of chronic GVHD changes. Our results also confirm the need for clinical trials to improve outcomes in patients with severe chronic GVHD, since their risk of non-relapse mortality remains elevated compared to that of patients with mild or moderate GVHD.

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

1. Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. Biol Blood Marrow Transplant. 2003;9(4):215-33.
2. Socie G, Schmoor C, Bethge WA, Ottinger HD, Stelljes M, Zander AR, 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):6375-82.
3. Weiden PL, Sullivan KM, Flournoy N, Storb R, Thomas ED. Antileukemic effect of chronic graft-versus-host disease: contribution to improved survival after allogeneic marrow transplantation. N Engl J Med. 1981;304(25):1529-33.
4. Horowitz MM, Gale RP, Sondel PM, Goldman JM, Kersey J, Kolb HJ, et al. Graft-versus-leukemia reactions after bone marrow transplantation. Blood. 1990;75(3):555-62.
5. Kolb HJ, Schmid C, Barrett AJ, Schendel DJ. Graft-versus-leukemia reactions in allogeneic chimeras. Blood. 2004;103(3):767-76.
6. Thepot S, Zhou J, Perrot A, Robin M, Xhaard A, de Latour RP, et al. The graft-versus-leukemia effect is mainly restricted to NIH-defined chronic graft-versus-host disease after reduced intensity conditioning before allogeneic stem cell transplantation. Leukemia. 2010;24(11):1852-8.
7. Inamoto Y, Flowers ME, Lee SJ, Carpenter PA, Warren EH, Deeg HJ, et al. Influence of immunosuppressive treatment on risk of recurrent malignancy after allogeneic hematopoietic cell transplantation. Blood. 2011;118(2):456-63.
8. Filipovich AH, Weisdorf D, Pavletic S, Socie G, Wingard JR, Lee SJ, 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):945-56.
9. Lee SJ, Klein JP, Barrett AJ, Ringden O, Antin JH, Cahn JY, et al. Severity of chronic graft-versus-host disease: association with treatment-related mortality and relapse. Blood. 2002;100(2):406-14.
10. Perez-Simon JA, Encinas C, Silva F, Arcos MJ, Diez-Campelo M, Sanchez-Guijo FM, et al. Prognostic factors of chronic graft-versus-host disease following allogeneic peripheral blood stem cell transplantation: the National Institutes Health scale plus the type of onset can predict survival rates and the duration of immunosuppressive therapy. Biol Blood Marrow Transplant. 2008; 14(10):1163-71.
11. Cho BS, Min CK, Eom KS, Kim YJ, Kim HJ, Lee S, et al. Feasibility of NIH consensus criteria for chronic graft-versus-host disease. Leukemia. 2009;23(1):78-84.
12. Arai S, Jagasia M, Storer B, Chai X, Pidala J, Cutler C, et al. Global and organ-specific chronic graft-versus-host disease severity according to the 2005 NIH Consensus Criteria. Blood. 2011;118(15):4242-9.
13. Kuzmina Z, Eder S, Bohm A, Pernicka E, Vormittag L, Kalhs P, et al. Significantly worse survival of patients with NIH-defined chronic graft-versus-host disease and thrombocytopenia or progressive onset type: results of a prospective study. Leukemia. 2012;26(4):746-56.
14. 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):1114-20.
15. Wingard JR, Piantadosi S, Vogelsang GB, Farmer ER, Jabs DA, Levin LS, et al. Predictors of death from chronic graft-versus-host disease after bone marrow transplantation. Blood. 1989;74(4):1428-35.
16. Vigorito AC, Campregher PV, Storer BE, Carpenter PA, Moravec CK, Kiem HP, et al. Evaluation of NIH consensus criteria for classification of late acute and chronic GVHD. Blood. 2009;114(3):702-8.
17. Arora M, Klein JP, Weisdorf DJ, Hassebroek A, Flowers ME, Cutler CS, et al. Chronic GVHD risk score: a Center for International Blood and Marrow Transplant Research analysis. Blood. 2011;117(24):6714-20.
18. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr. 1974;19:716-23.
19. Singer JD, Willett JB. Describing Continuous-time Event Occurrence Data. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. New York: Oxford University Press, 2003:468.
20. Jacobsohn DA, Kurland BF, Pidala J, Inamoto Y, Chai X, Palmer JM, et al. Correlation between NIH composite skin score, patient reported skin score and outcome: results from the Chronic GVHD Consortium. Blood. 2012;120(13):2545-52.
21. Pidala J, Chai X, Kurland BF, Inamoto Y, Flowers ME, Palmer J, et al. Analysis of gastrointestinal and hepatic chronic graft-versus-host disease manifestations on major outcomes: a Chronic Grant-Versus-Host Disease Consortium study. Biol Blood Marrow Transplant. 2013;19(5):784-91.
22. Clark JG, Crawford SW, Madtes DK, Sullivan KM. Obstructive lung disease after allogeneic marrow transplantation. Clinical presentation and course. Ann Intern Med. 1989;111(5):368-76.
23. Dudek AZ, Mahaseth H, DeFor TE, Weisdorf DJ. Bronchiolitis obliterans in chronic graft-versus-host disease: analysis of risk factors and treatment outcomes. Biol Blood Marrow Transplant. 2003;9(10):657-66.
24. Bacigalupo A, Chien J, Barisione G, Pavletic S. Late pulmonary complications after allogeneic hematopoietic stem cell transplantation: diagnosis, monitoring, prevention, and treatment. Semin Hematol. 2012;49(1): 15-24.
25. Jacobs R, Tran U, Chen H, Kassim A, Engelhardt BG, Greer JP, et al. Prevalence and risk factors associated with development of ocular GVHD defined by NIH consensus criteria. Bone Marrow Transplant. 2012;47(11):1470-3.