



Published in final edited form as:

Bone Marrow Transplant. 2016 October ; 51(10): 1350–1353. doi:10.1038/bmt.2016.131.

Confounding Factors Affecting the National Institutes of Health (NIH) Chronic Graft-Versus-Host Disease Organ-Specific Score and Global Severity

Sahika Zeynep Aki^{1,3,*}, Yoshihiro Inamoto^{1,*}, Paul A. Carpenter^{1,2}, Barry E. Storer^{1,2}, Brenda M. Sandmaier^{1,2}, Stephanie J. Lee^{1,2}, Paul J. Martin^{1,2}, and Mary E.D. Flowers^{1,2}

¹Fred Hutchinson Cancer Research Center

²University of Washington, Seattle, WA, USA

³Gazi University Faculty of Medicine Department of Hematology, Ankara, Turkey

Abstract

The 2005 NIH chronic graft-versus-host disease (cGVHD) organ severity is based on the assessment of current status regardless of whether abnormalities are due to GVHD. The score assignment does not require knowledge of past manifestations, attribution, or whether cGVHD is still active. The aim of this study is to describe confounding factors affecting organ scores in patients with cGVHD. The study included 189 consecutive cGVHD patients evaluated at our center in 2013. Providers completed the NIH 0–3 organ-specific scoring evaluation with two questions added for each organ to identify abnormalities that were 1) not attributed to cGVHD, or 2) attributed to cGVHD plus other causes. Abnormalities attributed to causes other than GVHD were recorded. Eighty (14%) abnormalities were not attributed to cGVHD in at least one organ, and 41 (7%) abnormalities were attributed to cGVHD plus other causes in at least one organ. A total of 436 (78%) abnormalities were attributed only to cGVHD. Abnormalities not attributed to cGVHD were observed most frequently in the lung, gastrointestinal tract and skin. Most common abnormalities included pre-transplant condition, sequelae from GVHD, deconditioning, infections and medications. Our results support the the 2014 NIH consensus recommendation to consider attribution when scoring organ abnormalities.

Keywords

Chronic GVHD; NIH consensus criteria; allogeneic hematopoietic cell transplantation; confounder

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Address for Correspondence: Mary Flowers, MD, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue N., Mail Stop D5-290, PO Box 19024, Seattle, WA 98109. mflowers@fhcrc.org.

*These authors contributed equally to this manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

INTRODUCT ON

Chronic graft-versus-host disease (GVHD) is a major complication of allogeneic hematopoietic cell transplantation (HCT) with significant effects on quality of life, survival and other transplant outcomes. Chronic GVHD occurs in 30 % to 50 % of allogeneic HCT recipients and is often associated with significant late morbidity and mortality.¹⁻⁶ Chronic GVHD is an immune-mediated disorder affecting multiple organ systems with heterogeneous clinical manifestations. In 2005, the National Institutes of Health (NIH) Consensus Working Group for Diagnosis and Staging of Chronic Graft-versus-Host Disease proposed criteria for diagnosis, subcategorization of acute and chronic GVHD, organ-specific severity scoring, global severity assessment, and indications for systemic treatment.⁷ Organ-specific severity scoring was based on assessment of the current status of each organ or site without a requirement for knowledge of past manifestations, attribution of abnormalities to chronic GVHD, or belief that the manifestation reflected an active versus a fixed deficit; clinicians simply scored the patient's current condition.

Several prospective studies have reported that organ severity scoring per the 2005 NIH Chronic GVHD Consensus Criteria is a valid and sensitive method for evaluating individual organ involvement and clinical monitoring.⁸⁻¹² Individual organ scoring may be combined as an NIH global severity score that showed marked associations with functional, quality of life, and survival outcomes measures.¹³⁻¹⁵ Despite these standardized criteria for chronic GVHD organ scoring and global severity, their application in clinical practice raised many questions about guidance to score abnormalities without considering past manifestations, fixed deficits, or whether abnormalities are attributable to chronic GVHD or another cause. Two key practical questions about the 2005 NIH organ scoring consensus criteria are: 1) Should we score symptoms or signs not attributed to chronic GVHD and 2) if yes, how should these organ scores be incorporated into the calculation of global severity. The aim of this prospective cross-sectional study was to describe the confounding factors affecting the 2005 NIH 0-3 organ score and global severity in patients with chronic GVHD.

MATERIALS AND METHODS

Patients

All outpatients were prospectively evaluated for GVHD manifestations using a standardized assessment form between January and December 2013 at the Long-Term Follow-Up (LTFU) clinic of Seattle Cancer Care Alliance/Fred Hutchinson Cancer Research Center. Information from patient visits that met all of the following was used for the analysis: (1) the first visit from each patient during the study period, and (2) diagnosis of chronic GVHD (*classic* or *overlap* subtypes) within 3 years after HCT according to the NIH consensus criteria. Patients gave written informed consent allowing their medical records to be used for research, in accordance with the Declaration of Helsinki, and the Institutional Review Board of the Fred Hutchinson Cancer Research Center approved the study.

GVHD assessment form

The standard assessment form for organ severity recommended by the 2005 NIH consensus criteria was modified to include two additional questions for each organ site to identify abnormalities that were 1) not attributed to chronic GVHD (yes/no), or 2) attributed to chronic GVHD plus other causes (yes/no).⁷ If abnormalities not attributed to chronic GVHD were present, causes of the abnormalities were documented by the medical evaluator. For example a patient with elevated liver enzyme tests that may be due to chronic GVHD or medication would be reported under the category of “attributed to chronic GVHD plus other causes.” In addition to capturing organ score plus the attribution of the abnormalities, information about abnormalities in male genitalia and about presence or absence of diagnostic or distinctive features of the mouth in asymptomatic patients were also included in the assessment form.

Statistical analysis

Descriptive statistics were used to summarize the patient-, transplantation- and GVHD-related characteristics. Continuous variables are presented as median and range, and categorical variables are presented as frequency and percentage. Global severity of chronic GVHD (none, mild, moderate or severe) was calculated from individual organ scores using the NIH consensus criteria algorithm. The higher value was used for final scoring of the lung if discrepancy existed between pulmonary symptom score and pulmonary function test (PFT) scores as recommended in the 2005 NIH consensus criteria for diagnosis and staging of chronic GVHD. Global severity was recalculated with the same algorithm after organ sites solely with abnormalities not attributed to chronic GVHD were scored as zero. In this analysis, the scores of sites that were “attributable to chronic GVHD plus other causes” were not changed.

RESULTS

Patient and transplantation characteristics

A total of 189 patients with chronic GVHD were evaluated by the LTFU clinical service between January 2013 and December 2013. Patient and transplantation characteristics are shown in Table 1. The median patient age at transplantation was 51 (range, 9 to 75) years. Nearly all patients (97%) were adults (greater than 18 years of age). Of the 189 patients, 158 (84%) received mobilized blood cells for transplantation, 110 (58%) had an HLA-matched unrelated donor, and 99 (53%) received myeloablative conditioning. The median interval from transplantation to chronic GVHD diagnosis was 6.9 (range, 2.4 to 37.6) months. The median interval from diagnosis of chronic GVHD to assessment for this study was 8.7 (range, 0 to 39.2) months.

The sites most frequently involved at the time of assessment were skin (59%), mouth (53%), eyes (53%), and lungs (43%) (Figure 1). Sixty out of the 189 patients (32%) presented with diagnostic or distinctive oral features on examination but without any oral symptoms.. One hundred sixty four patients (87%) had available PFT results to allow lung scoring.

Among the 189 patients, 436 (78%) abnormalities were scored and attributed only to chronic GVHD. Eighty (14%) abnormalities were not attributed to chronic GVHD in at least 1 organ, and 41 (7%) of abnormalities were attributed to chronic GVHD plus other causes in at least 1 organ (Figure 1A).

Abnormality and its attribution at individual organ sites

Abnormalities not attributed to chronic GVHD—Proportions of abnormalities not attributed to GVHD differed according to organ sites, ranging from 2% to 39% (Figure 1). Gastrointestinal tract (39%), lungs (28%), skin (15%), and joints (15%) were the organ sites most frequently affected by abnormalities not attributed to chronic GVHD. Causes of abnormalities not attributed to chronic GVHD are summarized in Table 2. Overall, the most frequent causes were abnormalities present prior to transplantation (36%), followed by medication effects (18%), deconditioning (18%), infections (16%), and sequelae from prior GVHD skin involvement (i.e., hyperpigmentation) (9%). Deconditioning refers to tiredness and/or dyspnea only on exertion that is not associated with abnormal pulmonary function tests nor explained by other causes in patients who spent most of the day sitting or lying down (sedentary) who endorse being “out of shape”. All manifestations reported as sequelae in the skin were hyperpigmentation. Gastrointestinal abnormalities related to side effects of treatment were diarrhea caused by oral magnesium supplementation and constipation caused by narcotics. Pulmonary abnormalities were caused by deconditioning mostly due to prolonged use of corticosteroids and physical inactivity. These patients had stable mild shortness of breath with normal or stable PFTs.

Abnormalities attributed to chronic GVHD plus other causes—In 2% to 19% of organ scorings symptoms or signs were reported as abnormalities attributed to chronic GVHD plus other causes (Figure 1). For these patients, the liver and gastrointestinal tract were the most prevalent organs (19% and 16%, respectively) and medication side effects, infections, iron overload, and fatty liver were the reported contributing factors for abnormalities present in addition to GVHD.

NIH global severity with or without consideration of confounders

Chronic GVHD global severity according to the 2005 NIH algorithm without consideration of attribution was mild in 35 patients (19%), moderate in 109 patients (58%), and severe in 45 patients (24%). After organs and sites with causes not attributed solely to chronic GVHD (n=62) were re-scored as zero, recalculation led to down-grading of the global severity in 18 patients (9.5%) (Table 3). Overall severity of chronic GVHD with and without confounders is shown in Figure 2. Four patients with moderate or severe global severity had no chronic GVHD when confounders were considered at the scoring evaluation. Manifestations attributed to scoring in these patients included dry skin (n=2), chronic obstructive pulmonary disease (n=1), irradiation pulmonary fibrosis (n=1), deconditioning as the cause of exertional dyspnea (n=2), infectious conjunctivitis (n=1) and nausea related to dasatinib (n=1). Immunosuppressive treatment was tapered in all patients.

DISCUSSION

The NIH consensus criteria provides standardized tools for reporting chronic GVHD (7). While many aspects of the 2005 NIH consensus criteria have been validated by several retrospective and prospective studies,^{8–18} daily practice has raised questions about the clinical scoring of organs.¹⁹ For instance, the rules for clinically scoring organs do not stipulate whether the abnormality is due to chronic GVHD or other causes. Further, the lack of distinction between active disease (e.g., erythematous rash) and fixed deficits or sequelae (e.g., hyperpigmentation as a consequence from prior rash) provides additional controversy in the implementation of the 2005 NIH consensus criteria.⁷

In this prospective study, we first examined the incidence and causes of abnormalities not attributed to chronic GVHD in each organ, and then recalculated the global severity by assigning an organ score of zero when abnormalities were not attributed to chronic GVHD. We have demonstrated that overall 14% of abnormal organ scores were attributed only to causes other than GVHD, and this was especially apparent in the lung (28%) and liver (39%). In descending order of importance pre-transplant conditions, medication effects, deconditioning and infection accounted for the majority of abnormal NIH scores unattributed to GVHD. One exception was the skin, where hyperpigmentation, the sequela of active rash, was equally important as a pre-transplant condition in accounting for abnormal skin scores that were not attributed to GVHD. We also observed that 2% to 19% of abnormal organ scores were attributed to chronic GVHD plus other causes. These observations suggest that failure to consider the attribution of organ score abnormalities might result in inappropriately high global severity scores. In our study, near 10% of global severity scores would be lower if certain manifestations were excluded because of these considerations. Four patients with moderate or severe global severity had no chronic GVHD when confounders were considered at the scoring evaluation, suggesting that physicians consider the causes of abnormalities when they make decisions regarding the dose and duration of immunosuppressive treatment.

In a recent survey of controversial areas in the 2005 NIH consensus criteria for diagnosis and scoring of chronic GVHD, investigators showed high agreement on the following two topics. First, signs, symptoms and diagnostic test results clearly not due to chronic GVHD should be scored differently. Second, active disease and fixed deficits should be distinguished in organ severity scoring.¹⁹ Our results support this unmet need from the real-world experiences. Notably, a third of patients with chronic GVHD had organ abnormalities from causes other than GVHD. In clinical conditions where physical exam may be insufficient to distinguish active disease and fixed deficits, identification of biomarkers or molecular studies of biopsies may potentially be useful in differentiating these processes.

The NIH chronic GVHD consensus criteria has been recently revised.²⁰ The current consensus incorporates asymptomatic organ manifestations (e.g., asymptomatic oral chronic GVHD). Our results support this change since 32% of HCT recipients presented with asymptomatic diagnostic or distinctive features of mouth involvement.

This study has several limitations. First, this is a single center descriptive study using a relatively small cohort. Correlation with outcomes according to different scoring algorithms needs to be examined in future studies. Second, most of the patients were adults and the results might not apply to children. Third, although medical records were checked after data entry regarding the organ and site specific abnormalities not attributed to chronic GVHD, the assessment of chronic GVHD and attribution of abnormalities were done by different medical providers who differed in their chronic GVHD expertise. Lastly, quantifying the extent of attribution to chronic GVHD versus clinical confounder remains an open question in the field.

Our results indicate that an unmodified 2005 NIH consensus organ scoring form will collect a substantial proportion of abnormalities attributed to causes other than chronic GVHD. Consideration of “attribution” confounders yields a slight downward shift in the distribution of global severity scores, thereby correcting an overestimation of organ involvement by chronic GVHD. This shift could have consequences for treatment decisions since the NIH consensus criteria advise systemic treatment only for moderate to severe chronic GVHD.⁷ Our results support the changes recommended by the new 2014 NIH consensus criteria.²⁰

Acknowledgments

This work was supported by grants CA018029, CA118953, CA078902 (National Cancer Institute) and HL122173 (National Heart, Lung, and Blood Institute) at the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

REFERENCES

1. Bashey A, Zhang X, Sizemore CA, Manion K, Brown S, Holland HK, et al. T-cell-replete HLA-haploidentical hematopoietic transplantation for hematologic malignancies using post-transplantation cyclophosphamide results in outcomes equivalent to those of contemporaneous HLA-matched related and unrelated donor transplantation. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2013; 31(10):1310–1316. [PubMed: 23423745]
2. Anasetti C, Logan BR, Lee SJ, Waller EK, Weisdorf DJ, Wingard JR, et al. Peripheral-blood stem cells versus bone marrow from unrelated donors. *The New England journal of medicine*. 2012; 367(16):1487–1496. [PubMed: 23075175]
3. Mielcarek M, Storer B, Martin PJ, Forman SJ, Negrin RS, Flowers ME, et al. Long-term outcomes after transplantation of HLA-identical related G-CSF-mobilized peripheral blood mononuclear cells versus bone marrow. *Blood*. 2012; 119(11):2675–2678. [PubMed: 22308289]
4. Saber W, Opie S, Rizzo JD, Zhang MJ, Horowitz MM, Schriber J. Outcomes after matched unrelated donor versus identical sibling hematopoietic cell transplantation in adults with acute myelogenous leukemia. *Blood*. 2012; 119(17):3908–3916. [PubMed: 22327226]
5. Alousi AM, Le-Rademacher J, Saliba RM, Appelbaum FR, Artz A, Benjamin J, et al. Who is the better donor for older hematopoietic transplant recipients: an older-aged sibling or a young, matched unrelated volunteer? *Blood*. 2013; 121(13):2567–2573. [PubMed: 23361908]
6. Dazzi F, Szydlo RM, Craddock C, Cross NC, Kaeda J, Chase A, et al. Comparison of single-dose and escalating-dose regimens of donor lymphocyte infusion for relapse after allografting for chronic myeloid leukemia. *Blood*. 2000; 95(1):67–71. [PubMed: 10607686]
7. 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. *Biology of Blood and Marrow Transplantation*. 2005; 11(12):945–956. [PubMed: 16338616]

8. Inamoto Y, Chai X, Kurland BF, Cutler C, Flowers ME, Palmer JM, et al. Validation of measurement scales in ocular graft-versus-host disease. *Ophthalmology*. 2012; 119(3):487–493. [PubMed: 22153706]
9. 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):2545–2552. [PubMed: 22773386]
10. Treister N, Chai X, Kurland B, Pavletic S, Weisdorf D, Pidala J, et al. Measurement of oral chronic GVHD: results from the Chronic GVHD Consortium. *Bone Marrow Transplantation*. 2013 [Epub ahead of print 2013 Jan 2028] doi: 2010.1038/bmt.2012.2285.
11. 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 Graft-Versus-Host Disease Consortium study. *Biology of Blood & Marrow Transplantation*. 2013; 19(5): 784–791. [PubMed: 23395601]
12. Palmer J, Williams K, Inamoto Y, Chai X, Martin PJ, Tomas LS, et al. Pulmonary symptoms measured by the National Institutes of Health lung score predict overall survival, nonrelapse mortality, and patient-reported outcomes in chronic graft-versus-host disease. *Biology of Blood and Marrow Transplantation*. 2014; 20(3):337–344. [PubMed: 24315845]
13. Pidala J, Kim J, Anasetti C, Nishihori T, Betts B, Field T, et al. The global severity of chronic graft-versus-host disease, determined by National Institutes of Health consensus criteria, is associated with overall survival and non-relapse mortality. *Haematologica*. 2011; 96(11):1678–1684. [PubMed: 21791465]
14. 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):4242–4249. [PubMed: 21791424]
15. Baird K, Steinberg SM, Grkovic L, Pulanic D, Cowen EW, Mitchell SA, 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. *Biology of Blood & Marrow Transplantation*. 2013; 19(4):632–639. [PubMed: 23340040]
16. Jagasia M, Chinratanalab W, Giglia J, Dixon S, Chen H, Frangoul H, et al. Incidence and outcome of chronic graft-versus-host disease (CGVHD) after allogeneic stem cell transplant (SCT) using National Institute of Health (NIH) consensus criteria. *Biology of Blood and Marrow Transplantation*. 2007; 13(Suppl. 2)(2):110–111. #303.
17. Arora M, Nagaraj S, Witte J, Defor TE, MacMillan M, Burns LJ, et al. New classification of chronic GVHD: added clarity from the consensus diagnoses. *Bone Marrow Transplantation*. 2009; 43(2):149–153. [PubMed: 18794869]
18. 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): 702–708. [PubMed: 19470693]
19. Inamoto Y, Jagasia M, Wood WA, Pidala J, Palmer J, Khera N, et al. Investigator feedback about the 2005 NIH diagnostic and scoring criteria for chronic GVHD. *Bone Marrow Transplantation*. 2014; 49(4):532–538. [PubMed: 24464142]
20. Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group Report. *Biology of Blood & Marrow Transplantation*. 2015; 21(3):389–401. [PubMed: 25529383]

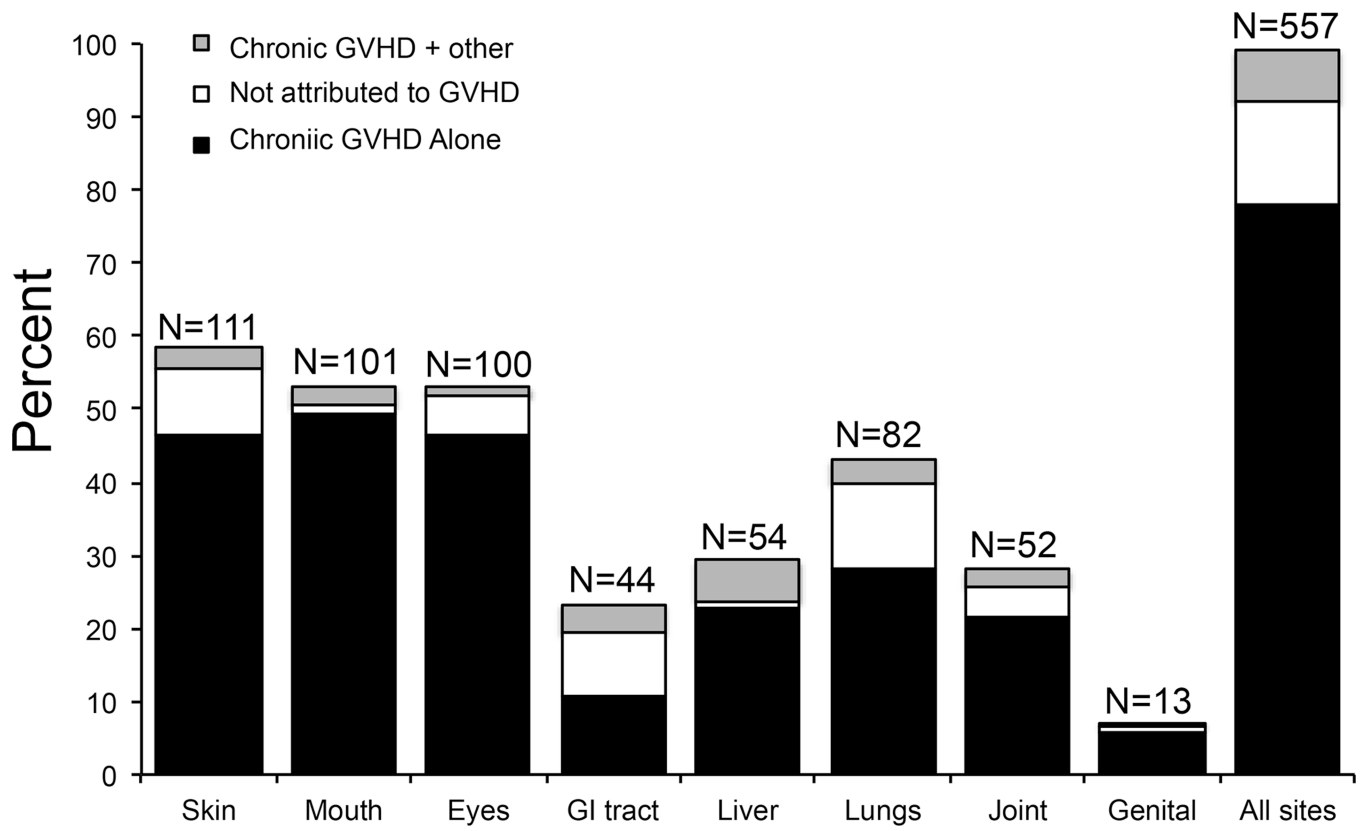


Figure 1. Percentage of abnormal NIH scores by organ and according to attribution.

Table 1

Patient and chronic GVHD characteristics

Characteristic	N = 189
Median patient age at transplantation, y (range)	51 (9 – 75)
Patient gender, no. (%)	
Male	127 (67)
Female	62 (33)
Donor-patient gender combination[*], no. (%)	
Female to male	60 (33)
Other	124 (67)
Diagnosis, no. (%)	
Myeloid malignancy	106 (56)
Lymphoid malignancy	74 (39)
Other/non-malignant	9 (5)
Conditioning regimen, no. (%)	
High intensity without high-dose TBI	64 (34)
High intensity with high-dose TBI	35 (19)
Reduced intensity	90 (48)
Graft source, no. (%)	
Bone marrow	19 (10)
Mobilized peripheral blood cells	158 (84)
Cord blood	12 (6)
HLA matching and donor type, no. (%)	
HLA-matched related	47 (25)
HLA-matched unrelated	110 (58)
HLA-antigen or allele-mismatched	32 (17)
Months from transplantation to chronic GVHD diagnosis, median (range)	6.9 (2.4 – 37.6)
Months from diagnosis of chronic GVHD to assessment, median (range)	8.7 (0 – 39.2)
Involved site, no. (%)	
Skin	111 (59)
Mouth	101 (53)
Eyes	100 (53)
Gastrointestinal tract	44 (23)
Liver	54 (29)
Lungs	82 (43)
Joint or fascia	52 (28)
Genital tract [†]	14 (7)
No. of sites involved, no. (%)	

Characteristic	N = 189
1 or 2	83 (44)
3	44 (23)
4	62 (33)
Subcategory of chronic GVHD	
Classic	83 (44)
Overlap	106 (56)

TBI, total body irradiation; HLA, human leukocyte antigen.

* Donor gender was unknown for 5 male patients.

[†] Unknown for 10 patients. Of the 14 patients with genital track abnormalities one was male.

Table 2

NIH Organ score abnormalities attributed to causes other than chronic GVHD

Cause, no. (%)	Skin	Mouth	Eyes	GI tract	Liver	Lungs	Joint	Genital	Total
Total, no.	17	2	10	17	2	23	8	1	80
Pretransplant condition	8 (47)	0 (0)	5 (50)	5 (29)	0 (0)	5 (22)	5 (63)	1 (100)	29 (36)
Medication effect	0 (0)	1 (50)	2 (20)	10 (59)	1 (50)	0 (0)	0 (0)	0 (0)	14 (18)
Deconditioning	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	12 (52)	2 (25)	0 (0)	14 (18)
Infections	2 (12)	1 (50)	2 (20)	2 (12)	0 (0)	6 (26)	0 (0)	0 (0)	13 (16)
Sequela	7 (41)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	7 (9)
Other	0 (0)	0 (0)	1 (10)	0 (0)	1 (50)	0 (0)	1 (13)	0 (0)	3 (4)

Table 3

NIH global severity with or without consideration of confounders

	Global severity without confounder, N (%)			
	None	Mild	Moderate	Severe
				Total
Global severity with confounder, N (%)				
Mild	4	31	0	0
Moderate	3	8	98	0
Severe	1	2	0	42
Total	8 (4)	41 (22)	98 (52)	42 (22)
				189 (100)