Predictive models for ocular chronic graft-*versus*-host disease diagnosis and disease activity in transplant clinical practice

Lauren M. Curtis,[1] Manuel B. Datiles III,[2] Seth M. Steinberg,[3] Sandra A. Mitchell,[4] Rachel J. Bishop,[2] Edward W. Cowen,[5] Jacqueline Mays,[6] John M. McCarty,[7] Zoya Kuzmina,[1] Filip Pirsl,[1] Daniel H. Fowler,[1] Ronald E. Gress[1] and Steven Z. Pavletic[1]

[1]Experimental Transplantation and Immunology Branch, National Cancer institute (NCI), National Institutes of Health (NIH), Bethesda, MD; [2]National Eye Institute, NIH, Bethesda, MD; [3]Biostatistics and Data Management Section, NCI, NIH, Rockville, MD; [4]Outcomes Research Branch, Division of Cancer Control and Population Sciences, NCI, NIH, Rockville, MD; [5]Dermatology Branch, NCI, NIH, Bethesda, MD; [6]National Institutes of Dental and Craniofacial Research, NIH, Bethesda, MD; and [7]Bone Marrow Transplant Program, Massey Cancer Center, Virginia Commonwealth University Medical Center, Richmond, VA, USA

ABSTRACT

Ocular chronic graft-*versus*-host disease is one of the most bothersome common complications following allogeneic hematopoietic stem cell transplantation. The National Institutes of Health Chronic Graft-*versus*-Host Disease Consensus Project provided expert recommendations for diagnosis and organ severity scoring. However, ocular chronic graft-*versus*-host disease can be diagnosed only after examination by an ophthalmologist. There are no currently accepted definitions of ocular chronic graft-*versus*-host disease activity. The goal of this study was to identify predictive models of diagnosis and activity for use in clinical transplant practice. A total of 210 patients with moderate or severe chronic graft-*versus*-host disease were enrolled in a prospective, cross-sectional, observational study (*clinicaltrials.gov identifier: 00092235*). Experienced ophthalmologists determined presence of ocular chronic graft-*versus*-host disease, diagnosis and activity. Measures gathered by the transplant clinician included Schirmer's tear test and National Institutes of Health 0-3 Eye Score. Patient-reported outcome measures were the ocular subscale of the Lee Chronic Graft-*versus*-Host Disease Symptom Scale and Chief Eye Symptom Intensity Score. Altogether, 157 (75%) patients were diagnosed with ocular chronic graft-*versus*-host disease; 133 of 157 patients (85%) had active disease. In a multivariable model, the National Institutes of Health Eye Score (P<0.0001) and Schirmer's tear test (P<0.0001) were independent predictors of ocular chronic graft-*versus*-host disease (sensitivity 93.0%, specificity 92.2%). The Lee ocular subscale was the strongest predictor of active ocular chronic graft-*versus*-host disease (P<0.0001) (sensitivity 68.5%, specificity 82.6%). Ophthalmology specialist measures that were most strongly predictive of diagnosis in a multivariate model were Oxford grand total staining (P<0.0001) and meibomian score (P=0.027). These results support the use of selected transplant clinician- and patient-reported outcome measures for ocular chronic graft-*versus*-host disease screening when providing care to allogeneic hematopoietic stem cell transplantation survivors with moderate to severe chronic graft-*versus*-host disease. Prospective studies are needed to determine if the Lee ocular subscale demonstrates adequate responsiveness as a disease activity outcome measure.

Introduction

Ocular chronic graft-*versus*-host disease (cGvHD) is a frequent long-term complication of allogeneic hematopoietic stem cell transplantation (HSCT), occurring in 40%-80% of patients.[1,2] Common symptoms are dryness, irritation, pain, redness, and blurred vision (Figure 1).[3] Corneal erosions and perforations can occur, and in rare cases of anterior chamber or choroid plexus involvement, blindness can result.[4,5]

The pathophysiology of ocular cGvHD is poorly understood, but is thought to be secondary to donor-derived T-cell mediated inflammatory processes. Due to homing signals, $CD4^+$ and $CD8^+$ T-lymphocytic infiltrates form in the periductal areas of lacrimal and meibomian glands, leading to accumulation of extracellular matrix (ECM), stromal $CD34^+$ fibroblasts and subsequent excessive fibrosis,[6,7] resulting in lacrimal gland function impairment. Tissue sections of the lacrimal glands have been found to exhibit increased expression of HSP47 in fibroblasts, which promotes excessive collagen accumulation.[8] Ocular surface abnormalities include infiltrates of T cells, with increased $CD8^+$: $CD4^+$ ratios[9] and increased expression of Th1-associated chemokines, such as IL-17.[10] Other pathways of inflammation leading to ocular cGvHD have been proposed, such as the upregulation of ICAM-1 expression in conjunctiva[11] and, more recently, the activation of the toll-like receptor 2 (TLR2) mediated NF-κB pathway.[12]

Unlike sclerotic skin cGvHD, which is a diagnostic manifestation of cGvHD, ocular cGvHD must be confirmed by an ophthalmologist.[13,14] In 2013, the International Ocular cGvHD Consensus Group published ophthalmology diagnostic guidelines for ocular cGvHD,[15] which recommend evaluation of the ocular surface disease index (OSDI), Schirmer's without anesthesia, corneal fluorescein staining and conjunctival injection, with each parameter assigned a severity score (0-3). Based on the aggregate score, and the presence of systemic

 doi:10.3324/haematol.2015.124131
The online version of this article has a Supplementary Appendix.
Manuscript received on January 21, 2015. Manuscript accepted on June 16, 2015.
Correspondence: pavletis@mail.nih.gov

cGvHD manifestations, a diagnosis of ocular cGvHD can be made. This diagnostic scoring system has not yet been prospectively validated, and exact ophthalmology criteria for cGvHD diagnosis remain to be defined.

The National Institutes of Health (NIH) cGvHD Consensus Project in 2005 proposed guidelines for transplant clinicians to assess ocular cGvHD severity; specified measures were the NIH eye score, Schirmer's tear test, Lee cGvHD symptom scoring and chief eye symptom score.[2,13] Apart from removing the Schirmer's tear test from the severity scoring scale, these recommendations did not change in 2014.[14] The ability of these criteria to predict ocular cGvHD diagnosis based on expert ophthalmologic examination is not known. In addition, there are no evidence-based recommendations regarding when transplant clinicians should suspect ocular cGvHD diagnosis and refer patients to an ophthalmologist for evaluation.

Accepted therapeutic response measures in ocular cGvHD are also lacking. The NIH cGvHD Consensus conference in 2005 proposed the Schirmer's tear test as a measure of response, but a large prospective longitudinal study by Inamoto *et al*. demonstrated that the Schirmer's tear test did not correlate well with either provider- or patient-reported perceptions of change in ocular cGvHD severity.[16] Additional investigations are needed to identify which specific measures are most strongly associated with ocular cGvHD activity, and thus might have been expected to have the best performance characteristics in measuring ocular cGvHD response in a clinical trial or in clinical practice.

Despite the high frequency of ocular cGvHD post HSCT, risk factors and clinical characteristics of ocular cGvHD are not completely understood. Matched related donor HSCT, male gender, prior acute skin GvHD, and oral and skin cGvHD involvement have been found to be associated with ocular cGvHD.[1,17] Reliable identification of risk factors for ocular cGvHD could select patients that might benefit from early intervention and provide insight into the pathophysiology of this condition.

The objective of this study is to determine which aspects of the NIH cGvHD severity scoring criteria are most predictive of ocular cGvHD diagnosis in a large, well-characterized cohort with moderate to severe cGvHD enrolled on a cross-sectional natural history study. Patients underwent a standardized ophthalmology specialist examination and each patient was determined as having ocular cGvHD diagnosis or not. Since there are no standard definitions for ocular cGvHD activity, a second objective of this study was to use ophthalmology expert decision as the gold standard (active *vs.* inactive ocular cGvHD) and determine which factors of the ophthalmology examination and transplant clinician examination correlate most closely with the presence of active ocular cGvHD.

Methods

The National Cancer Institute (NCI) cGvHD natural history protocol is a cross-sectional study in which patients present for a one-week multispecialty cGvHD assessment and clinical data collection.[18] This protocol was approved by the NCI Institutional Review Board (IRB) and all patients signed an IRB approved informed consent. Patients were examined by one of 2 experienced ophthalmologists (MBD, RJB) who assigned a severity score for each of the following: meibomian gland plugging; lid margin swelling, erythema and debris; conjunctival injection and conjunctival chemosis (Table 1). Visual acuity by Logmar,[19] tear film breakup time, Oxford staining grand total score (corneal and conjunctival) and Oxford corneal staining were also obtained.[20] Since two values were obtained for each patient (right eye, left eye), the

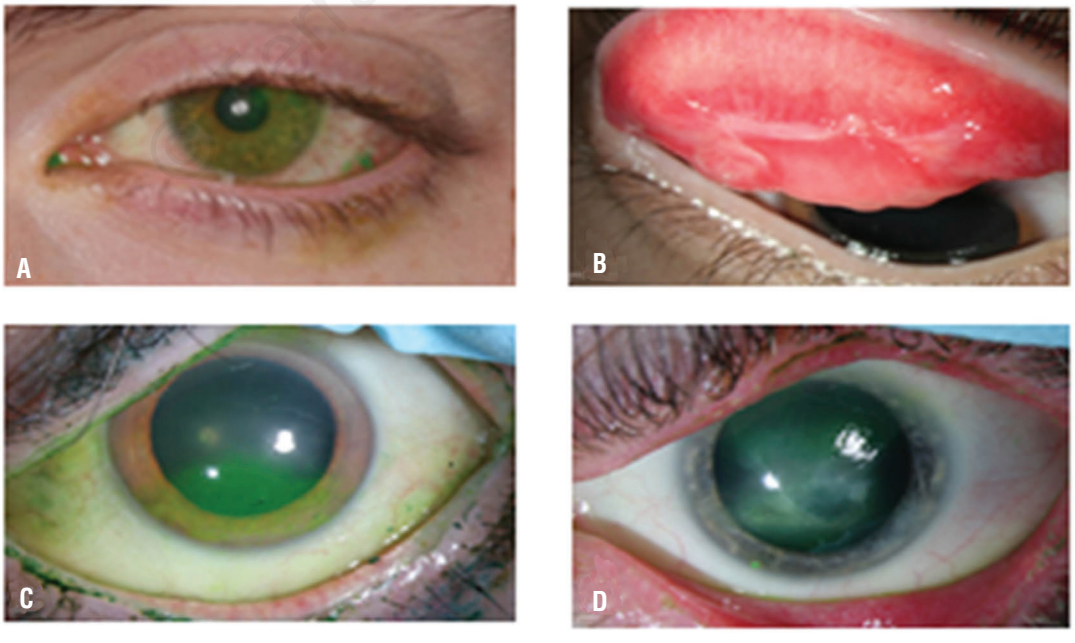

Figure 1. Images of ocular cGvHD. (A) Red, irritated, dry eyes. (B) Upper tarsal conjunctival scarring. (C) Punctate keratopathy. D) Corneal epithelial sloughing.

maximum value for Logmar, Oxford staining, meibomian score, lid swelling, lid erythema, tear film debris, conjunctival injection and chemosis was selected for each patient, and the minimum value selected for Schirmer's and tear film breakup time. Ophthalmologists confirmed the diagnosis of ocular cGvHD (yes-no) and classified it as active *versus* inactive based on their expert opinion.

In the transplant clinic, an eye dataset comprising of 4 measures was created for each patient (Table 2). Clinician-reported measures were the Schirmer's tear test without anesthesia and NIH 0-3 eye score.[13] Patient-reported measures were also obtained: chief eye symptom intensity score[21] and three items drawn from the Lee cGvHD Symptom Scale (bothered by dry eyes, needing to use eyedrops frequently, and difficulty seeing clearly).

Variables examined for ocular cGvHD risk were gender, age at study enrollment, cytomegalovirus (CMV) status of recipient, transplant conditioning (myeloablative *vs.* non-myeloablative), donor relationship (related *vs.* unrelated), donor gender (male *vs.* female), gender match (matched *vs.* unmatched), HLA match (matched *vs.* unmatched), type of GvHD prophylaxis used at transplant (tacrolimus-based, *vs.* cyclosporine-based, *vs.* T-cell depletion), receipt of donor lymphocyte infusion (DLI) post-transplant, prior acute graft-*versus*-host disease (aGvHD) of any type, prior aGvHD skin, aGvHD gastrointestinal (GI) tract, and/or aGvHD liver, and days from transplant to cGvHD diagnosis. CGvHD characteristics included: the type of onset (progressive, *de novo* or quiescent), cGvHD classification (classic, overlap or late acute), and cGvHD NIH organ severity scores for genital tract, joint/fascia, lungs, liver, GI tract, eyes, mouth, and skin. Measures were collected across the course of a 1-week cGvHD. Associations between ophthalmologist-, transplant clinician- and patient-reported measures and specialist confirmation of the diagnosis and degree of activity of cGvHD were examined through univariate analysis. The association between risk factors and

Table 1. Specialized ocular surface disease assessment by ophthalmologist.

Meibomian glands plugging (0-3) 0 = none; 1 = 1-2 glands; 2= 2-3 glands; 3 = all 5 glands, lid margin swelling
Lid margin swelling (0-4) 0 = normal; 1 = localized small region; 2 = diffuse, most lids; 3 = diffuse, most lids and protruding; 4 = diffuse and eversion of lids
Lid margin erythema (0-4) 0 = normal; 1= localized small region lids or skin; 2 = redness most lid margin or skin; 3 = redness most or all lid margin and skin; 4 = marked redness both lid margin and skin
Lid margin tear film debris (0-4) 0 = none; 1 = debris inferior tear meniscus; 2 = (1) and debris overlying cornea; 3 = (2) and presence of mucus; 4 = (3) and extensive mucus or filament
Conjunctival injection (0-4) 0 = normal; 1 = slight localized injection; 2 = pink palpebral or bulbar conjunctiva; 3 = red palpebral and/or bulbar conjunctiva; 4 = marked red palpebral and/or bulbar conjunctiva
Conjunctival chemosis (0-4) 0 = normal; 1 = slight localized swelling; 2 = moderate localized or mild diffuse swelling; 3 = severe diffuse swelling; 4 = prominent diffuse swelling

Table 2. Description of transplant clinician-reported and patient-reported ocular measures.

Schirmer's tear test (0-35 mm) Paper strips inserted into patient's eye (without anesthesia) and eyes closed for 5 minutes. Length of paper that has been moistened by tears is then measured in mm. Normal: ≥ 11 mm Mild-moderate: 6-10 mm Severe: 0-5 mm
NIH Eye score (0-3) 0: No symptoms 1. Mild dry eye symptoms not affecting activities of daily living (ADL) (requiring eyedrops <-3 times per day) OR asymptomatic signs of keratoconjunctivitis sicca 2. Moderate dry eye symptoms partially affecting ADL (requiring drops>3 times per day or punctual plugs) WITHOUT vision impairment 3. Severe dry eye symptoms significantly affecting ADL (special eyewear to relieve pain) OR unable to work because of ocular symptoms OR loss of vision caused by keratoconjunctivitis sicca
Chief eye symptom intensity score (0-10) Patient self-reporting of chief complaint with regard to eyes Rate how severe, in the past 7 days, is this eye symptom, between 0 (not at all severe), and 10 (most severe): 0 1 2 3 4 5 6 7 8 9 10
Lee cGvHD Symptom Scale Ocular Symptom items (0-4) Patient self-reporting of ocular symptoms (bothered by dry eyes, needing to use eyedrops frequently, and difficulty seeing clearly) in the past month: 0: not at all; 1: slightly; 2: moderately; 3: quite a bit; 4: extremely

Table 3. Demographic, clinical and transplant characteristics.

	N (%)
Total number of patients	210
Median age (range)	47 (10-70)
Gender	
Male	116 (55)
Female	94 (45)
Median days from transplant to enrollment (range)	1119 (124-8911)
Median days from cGvHD diagnosis to enrollment (range)	765 (20-6670)
Disease	
AML/ALL/MDS	96 (46)
HL/NHL/CLL/MM	71(34)
CML/IMF/MPD	22(10)
Other	21(10)
Conditioning regimen	
Myeloablative	112(53)
Non-myeloablative	95(45)
Unknown	3(1)
TBI conditioning	
Yes	79(38)
No	128(61)
Unknown	3(1)
GVHD prophylaxis	
Cyclosporine-based	83 (40)
Tacrolimus-based	80 (38)
T-cell depletion	33 (16)
Unknown	14(7)
Donor relationship	
Related	124(59)
HLA matched	112(53)
HLA mismatched	10(5)
Unknown	2(1)
Unrelated	84(40)
HLA Matched	55(26)
HLA mismatched	25(12)
Unknown	4(2)
Unknown	2(1)
Stem cell source	
Bone marrow	37(18)
Peripheral blood	166(79)
Cord	6(3)
Unknown	1(<1)
cGvHD onset[a]	
Progressive	73(35)
Quiescent	62(30)
de novo	72(35)
Unknown	3(1)
Lines of prior systemic therapy for cGvHD	
<2	21(10)
2-4	107(51)
>=5	81(39)
Unknown	1 (<1)
NIH global score[b]	
Mild	1 (<1)
Moderate	56(27)
Severe	153(73)
Intensity of current immunosuppression[c]	
None	40(19)
Mild	9(4)
Moderate	76(36)
High	83(40)
Unknown	2 (1)
cGvHD organ involvement	
Skin	167(80)
Mouth	135(64)
Eyes	166(79)
GI tract	95(45)
Liver	98(47)
Lungs	163(78)
Joints and fascia	134(64)
Genital (female only, n = 94)	54(57)
Active ocular cGvHD present	157(75)
	133(63)

%: calculated based on total number of patient (n=210), not taking into account missing data. AML: acute myeloid leukemia; ALL: acute lymphoblastic leukemia; MDS: myelodysplastic syndrome; HL: Hodgkin lymphoma; NHL; non-Hodgkin lymphoma; CLL: chronic lymphocytic leukemia; MM: multiple myeloma; CML: chronic myeloid leukemia; IMF: idiopathic myelofibrosis; MPD: myeloproliferative disorder; TBI: total body irradiation; HLA: human leukocyte antigen. [a]Definition for cGvHD onset are as follows: progressive (acute GvHD progressed directly to cGvHD); quiescent (acute GvHD resolved, then chronic GvHD developed); de novo (acute GvHD never developed). [b]Definition for NIH Global score as follows: mild (1 to 2 organs affected by cGvHD with scores 1); moderate (more than 2 organs with score 1, any score 2, or lung score 1), or severe (any score of 3 or lung score of 2).[13] [c]Definition of intensity of immunosuppresion is as follows: mild (single-agent prednisone 0.5 mg/kg/day); moderate (single agent prednisone 0.5 mg/kg/day and/or any single agent/modality); high (2 or more agents/modalities +/- prednisone 0.5 mg/kg/day).[29]

cGvHD characteristics and specialist-confirmed cGvHD diagnosis was also assessed using univariate statistics.

Between groups comparisons were performed using a Wilcoxon rank sum test (continuous parameters), Cochran-Armitage test for trend (ordered categorical parameters,[22] Fisher's exact test (dichotomous parameters), or Mehta's modification to Fisher's exact test (unordered categorical parameters).[23] Logistic regression modeling using step-wise elimination identified factors jointly associated with ophthalmologist-confirmed presence of ocular cGvHD, and factors jointly associated with active *versus* non-active cGvHD. All *P*-values are two-tailed and reported without any formal adjustment for multiple comparisons. In view of the large number of exploratory analyses performed, only univariate tests with $P<0.005$ were considered statistically significant, while those for which $0.005<P<0.05$ were interpreted as reflecting strong trends.

Results

Sample

A total of 293 adult and pediatric patients were enrolled between 2004 and 2013. The first 49 patients did not undergo a comprehensive eye examination with all ophthalmology measures, except for patient #26, since these standardized procedures had not yet been implemented in the protocol. Of the remaining patients (patient #50 onwards), 12 patients were removed from the analysis because the examining ophthalmologist was unable to discern the presence of ocular cGvHD, most often due to the presence of concurrent pathology such as infectious conjunctivitis. Seven patients were removed from the analysis because they were not found to have cGvHD in any organ system, including the eye. Eight patients were removed from the study because there were too many data missing from the ophthalmology evaluations, for reasons that included participant withdrawal from the study due to acute illness, lack of time to complete the eye examination due to scheduling conflicts or because an expert ophthalmologist was not available. Eight pediatric patients also did not undergo a full ophthalmology evaluation or complete the patient reported measures. The remaining cohort consisted of 210 patients whose data were used for the final analysis.

Demographics and transplant characteristics

Table 3 shows the demographic and transplant characteristics of the 210 patients. Median age was 47 years (range: 10-70 years); 55% were male (n=116) and 45% female (n=94). A total of 167 patients (80%) underwent an HLA matched allogeneic transplant, the majority receiving peripheral blood stem cells (n=166, 79%). The median time from transplant was 36.8 months (range: 4.1-292.8 months) and median time from cGvHD diagnosis to enrollment was 25.1 months (range: 0.7-219.1 months). The median number of lines of prior systemic therapy for cGvHD was 4 (range: 0-9). The majority of patients were judged to have clinically severe cGvHD based on the NIH Global score (n=153, 73%). A total of 157 patients (75%) were diagnosed with ocular cGvHD by an ophthalmologist, and of patients diagnosed with ocular cGvHD, 85% (133 of 157) were defined as having active ocular cGvHD.

Ophthalmologist-, transplant clinician- and patient-reported measures

The distribution of ophthalmologist measures and their association with the diagnosis of ocular cGvHD and "active" ocular cGvHD by univariate analysis is shown in Table 4. As expected, measures that identify and grade conjunctival, corneal, meibomian, and lid abnormalities were found to be the most strongly associated with the

Table 4. Ophthalmologist-reported measures and association with the diagnosis and activity level of ocular cGVHD.

	Ocular cGvHD Present n=157	Ocular cGvHD Absent n=53	P	Active ocular cGvHD n=133	Non-active ocular cGvHD n=24	P
Visual acuity						
Logmar			0.0095			0.0076
-0.3 to -0.01	3	2		2	1	
0	39	22		30	9	
>0	113	29		99	14	
Tear production						
Tear film breakup			0.010			0.0065
≤5	78	18		71	7	
>5	58	32		45	13	
Ocular surface staining						
Oxford grand total staining			<0.0001			<0.0001
0	7	22		3	4	
1-10	115	23		99	16	
11-15	18	0		18	0	
Oxford corneal staining			<0.0001			0.0002
0	22	34		15	7	
1-3	97	17		83	14	
4-5	32	0		31	1	
Specialized Ocular Surface Exam						
Meibomian score			<0.0001			0.0003
0	60	46		44	16	
1-2	79	3		75	4	
3-4	0	0		0	0	
Lid margin swelling			<0.0001			0.0004
0	63	44		45	18	
1-2	72	5		69	3	
3-4	5	0		5	0	
Lid margin erythema			<0.0001			0.0003
0	66	46		48	18	
1-2	72	3		69	3	
3-4	0	0		0	0	
Lid margin tear film debris			<0.0001			<0.0001
0	35	31		21	14	
1-2	83	18		76	7	
3-4	24	0		24	0	
Conjunctival injection score			<0.0001			<0.0001
0	52	42		35	17	
1-2	87	7		82	5	
3-4	8	0		8	0	
Conjunctival chemosis score			<0.0001			<0.0001
0	53	42		37	16	
1-2	85	7		80	5	
3-4	5	0		5	0	

Note: P-values for visual acuity, tear production, and Oxford grand total score determined based on continuous data; categorical values reported for consistency of presentation.

diagnosis of ocular cGvHD: Oxford grand total and corneal staining (both $P<0.0001$), meibomian score ($P<0.0001$), lid margin swelling ($P<0.0001$), lid margin erythema ($P<0.0001$), lid margin tear film debris ($P<0.0001$), conjunctiva injection score ($P<0.0001$), and conjunctiva chemosis score ($P<0.0001$). A multivariate analysis performed to determine which measures were most strongly predictive of ocular cGvHD diagnosis retained Oxford grand total ($P<0.0001$) and meibomian score ($P=0.027$). The ophthalmologist measures most strongly associated with "active" ocular cGvHD were the Oxford grand total ($P<0.0001$), Oxford corneal staining ($P=0.0002$), meibomian gland plugging ($P<0.0003$), lid margin swelling ($P=0.0004$), erythema ($P=0.0003$), tear margin debris ($P<0.0001$) and conjunctival injection ($P<0.0001$) and chemosis ($P<0.0001$). For a multivariable model predictive of active ocular cGvHD, the only ophthalmology measure retained was the Oxford grand total score ($P<0.0001$), which incorporates the severity staining score of both the cornea and conjunctiva.

The clinician- and patient-reported measures associated with the diagnosis of ocular cGvHD and active ocular cGvHD based on expert ophthalmology exam are seen in Table 5. The Schirmer's test ($P<0.0001$), item scores for

Table 5. Transplant clinician-reported measures and their association with the diagnosis and activity level of ocular cGVHD.

	Ocular cGvHD Present N=157	Ocular cGvHD Absent N=53	P	Active ocular cGvHD N=133	Non-active ocular cGvHD N=24	P
Schirmer's w/o anesthesia			<0.0001			0.004
≤5	142	12		122	20	
6-10	10	30		7	3	
11-19	3	7		2	1	
≥20	1	3		1	0	
Items from Lee cGVHD symptom scale						
Bothered by dry eyes			<0.0001			<0.0001
0	14	22		6	8	
1	18	17		14	4	
2	28	5		21	7	
3	41	5		39	2	
4	52	0		50	2	
Bothered by needing to use eyedrops frequently			<0.0001			<0.0001
0	23	28		10	13	
1	24	13		23	1	
2	18	7		13	5	
3	31	1		30	1	
4	57	0		54	3	
Bothered by difficulty seeing clearly			0.0001			0.0047
0	28	22		18	10	
1	31	10		27	4	
2	29	8		27	2	
3	38	7		33	5	
4	24	2		23	1	
NIH eye score			<0.0001			<0.0001
0	5	38		1	4	
1	59	14		43	16	
2	72	0		70	2	
3	21	0		19	2	
Chief eye symptom score			<0.0001			0.0023
0	2	14		0	2	
1	5	1		3	2	
2	7	4		5	2	
3	9	5		7	2	
4	13	2		11	2	
5	15	4		12	3	
6	11	3		9	2	
7	19	2		17	2	
8	20	2		19	1	
9	6	0		6	0	
10	12	0		11	1	

Schirmer's tear test and chief eye symptom score P-values were based on continuous representations; shown here as categorical values for consistency of presentation. P<0.005 were considered highly significant associations with ocular cGvHD and "active" ocular cGvHD, while 0.005 < P<0.05 reflect strong trends.

three items drawn from the Lee cGvHD Symptom Scale (bothered by dry eyes, needing to use eyedrops frequently, and difficulty seeing clearly) (all $P \leq 0.0001$), the NIH eye score ($P<0.0001$), and the chief eye symptom score ($P<0.0001$) were strongly associated with the diagnosis of ocular cGvHD. Among these measures, being bothered by dry eyes ($P<0.0001$), needing to use eyedrops frequently ($P<0.0001$), and the NIH eye score ($P<0.0001$) were the most strongly associated with active ocular cGvHD. Of the 5 clinician- and patient-reported measures included in this model, the NIH eye score (0 *vs.* 1 *vs.* 2 *vs.* 3) ($P<0.0001$) and lower Schirmer's tear test without anesthesia ($P<0.0001$) were significant independent predictors of the presence of ocular cGvHD. Table 6 shows the multivariable logistic models predictive of the diagnosis of ocular cGvHD. The resulting classification rule [2.9475 x NIH eye score – 0.2691 x Schirmer's (mm); if => 1.0986 then ocular cGvHD is present] correctly identified 93.0% of patients with ocular cGvHD and 92.2% of patients without ocular cGvHD based on the same patients used to create the models. In a model which included all 5 clinician- and patient-reported measures as covariates, the single item asking about being quite a bit or extremely bothered by dry eyes was the strongest independent predictor of ophthalmologist-diagnosed ocular cGvHD activity. This classification rule predicted active ocular cGvHD with 68.5% sensitivity and 82.6% specificity based on data used to create the model.

Risk factors and cGvHD characteristics

Related donor HSCT was associated with increased diagnoses of ocular cGvHD [103 of 124 (83%) of patients with related donors have ocular cGvHD *vs.* 54 of 84 (64%) of patients with unrelated donors; $P=0.0029$]. We also found that HLA matched HSCT was also a risk factor for ocular cGvHD, as 132 of 167 (79%) of patients with HLA matched HSCT are diagnosed with ocular cGvHD *versus* 20 of 35 (57%) with HLA mismatched HSCT ($P=0.0095$). Oral cGvHD was the only organ found to be strongly associated with ocular cGvHD ($P<0.0001$). None of the other parameters examined were found to be risk factors for ocular cGvHD, including type of GvHD prophylaxis (cyclosporine-based *vs.* tacrolimus-based *vs.* T-cell depleting) (*Online Supplementary Appendix*).

Discussion

The NIH cGvHD Consensus Project provides a set of recommendations and measures to define ocular cGvHD diagnosis and severity; however, with the exception of the Schirmer's tear test, these criteria have not been validated and tested against the gold standard for diagnosis, which is ophthalmology subspecialist assessment.[1,13] In addition, there is no accepted measure of ocular cGvHD activity and criteria for ocular cGvHD diagnosis by an ophthalmologist have still not been prospectively validated.[16]

In this analysis, we determined which of the ophthalmologist standard exam measures were most strongly associated with the diagnosis of ocular cGvHD, and found that the majority of the measures which graded abnormalities of the conjunctiva, cornea, meibomian glands, and lids were strongly associated with the presence of ocular cGvHD and active ocular cGvHD. In the multivariate analysis, however, two key ophthalmology measures were the most strongly predictive of ocular cGvHD diagnosis [higher Oxford grand total staining ($P<0.0001$) and higher meibomian gland score ($P=0.027$)], and identified higher Oxford grand total staining as most predictive of active ocular cGvHD ($P<0.0001$). Although it is generally recommended that a specialized eye examination include measurement of the tear film breakup time, in this analysis which included patients with moderate and severe cGvHD it was found to be less strongly associated with ocular cGvHD and active ocular cGvHD ($P=0.01$ and $P=0.007$, respectfuly). As this was a single center study, one of 2 ophthalmologists performed all examinations; our observations should be confirmed in a multicenter, longitudinal study, In addition, since the "gold standard" in this analysis was the presence of ocular cGvHD and active ocular cGvHD based on the opinion of the examining ophthalmologist, additional research is needed to examine the inter- and intra-rater reliability with which active cGvHD is identified and graded by ophthalmologists.

In a multivariable predictive model, higher NIH eye scores (3 *vs.* 2 *vs.* 1) ($P<0.0001$) and lower Schirmer's tear test values ($P<0.0001$) were significant independent predictors of ocular cGvHD (sensitivity 93.0%, specificity 92.2%). Therefore, this predictive model may identify which patients have a high likelihood of having ocular

Table 6. Multivariable logistic models predictive of the diagnosis of ocular cGvHD and active ocular cGvHD based on clinician- and patient-reported measures.

Predictive category	Type of measure	P	Classification rule	Sensitivity (%)	Specificity (%)
Diagnosis of ocular cGvHD	NIH eye	<0.0001	2.9475 x NIH Eye score - 0.2691 x Schirmer's tear test mm of wetting* If ≥ 1.0986 then ocular cGvHD is present	93.0	92.2
	Schirmer's w/o anesthesia	<0.0001	If < 1.0986, then ocular cGvHD is absent		
Active ocular cGvHD	Self-report of being bothered by dry eyes	<0.0001	If Lee dry eyes >2 then predict active ocular cGvHD If Lee dry eyes <=2 then predict NO active ocular cGvHD	68.5	82.6

*NIH: National Institutes of Health; w/o: without; Tot: total; cGvHD: chronic graft-versus-host disease. *Worst of the two eyes, without anesthesia.*

cGvHD in transplant clinical practice, and distinguishes those patients who may benefit from referral to ophthalmology for a comprehensive evaluation and treatment plan.[2] This multivariate model is a major improvement over the Schirmer's alone, which has previously been reported to predict ocular cGvHD with only 69% sensitivity and 58% specificity.[1] The latest 2014 NIH Consensus cGvHD guidelines suggest that the Schirmer's tear test no longer be obtained by a transplant clinician in order to determine ocular cGvHD severity due to its poor correlation with patient and clinician perceptions of change in severity.[14] However, we suggest that, in light of it being a significant independent predictor of presence of ocular cGvHD in this analysis, obtaining this test at the time of referral to an ophthalmologist for further evaluation is a reasonable approach. This recommendation is also consistent with the current recommendation for cGvHD diagnosis both by the NIH Consensus guidelines and by ophthalmology expert recommendations.[2,15]

Secondly, we identified measures that were found to be associated with the presence of active ocular cGvHD. Notably, the Schirmer's test is not as strongly associated with active ocular cGvHD compared to other transplant clinic measures, such as the Lee dry eye and Lee needs eyedrops frequently scores. This is in agreement with a prospective study by Inamoto *et al*. which demonstrated that the Schirmer's test is a suboptimal measure when evaluating change in severity of ocular cGvHD over time or in response to treatment.[16] Loss of tear production is likely to a permanent sequelae of T-cell infiltration and subsequent fibrosis in the periductal area of lacrimal ducts, changes which have been shown to be more prominent in patients with severe dry eye compared to mild dry eye.[6,7] Instead, a single item assessing being bothered by dry eye symptom from the patient's perspective (Lee dry eye score item >2) had 82.6% specificity in predicting the results of examination by an expert ophthalmologist. Although subjective and potentially modified by intervention in non-blinded studies, symptoms are at the center of the cGvHD patient's experience and validated symptom scales such as Lee scale could be a valuable and practical adjunct in assessing the disease activity in trials and clinical care.[24] The Lee eye subscale has previously been shown to be sensitive to change, and was recommended as a response measure in the up-dated NIH cGvHD Consensus Guidelines.[25] The sensitivity of the Lee dry eye score was, however, low for the predictive model presented here (68.5%), thus prospective studies are needed to determine if this single patient-reported measure assessing dry eye symptoms is sufficient for use as a clinical trial outcome and to guide therapeutic decision-making in clinical practice.

Thirdly, we identified an association between the HLA matched and related donor HSCT and the diagnosis of ocular cGvHD. Westenberg *et al*.[17] also reported the trend of increased risk of ocular cGvHD in patients with related donor allogeneic transplants *versus* unrelated donor transplants (60% *vs.* 45%; OR: 1.774; 95%CI: 0.801-3.929; *P*=0.166). In this current study, the relationship between related and HLA matched transplants and increased risk of ocular cGvHD was not explained by the type of GvHD prophylaxis received (tacrolimus-based *vs.* cyclosporine based vs. T-cell depleting). Although paradoxical, the association between an HLA matched HSCT and the risk of a particular cGvHD organ manifestation is not unique to ocular cGvHD. In a report on patients with sclerotic cGvHD, Inamoto *et al*. reported an increase in the incidence of sclerotic skin cGvHD in patients with HLA matched donors.[26] Future studies on the pathophysiology of cGvHD may shed light on these findings.

When examining the association between ocular cGvHD and other cGvHD characteristics, we found that oral cGvHD was associated with the diagnosis of ocular cGvHD. This was previously reported in an earlier, subgroup analysis of this patient cohort[27] in which a significant association between dry eye and dry mouth symptoms was detected, and also in other patient cohorts.[17] One proposed explanation for this association is the common developmental origins of the two organs, with involvement of ductal area target sites such as meibomian glands, lacrimal glands and salivary glands where similar infiltration patterns have been shown for T cells, fibroblasts, and other inflammatory cells.[7,15,28]

There are some limitations to this study. The cross-sectional nature of the study did not allow for multiple assessments over time. Therefore, determination of active ocular cGvHD was based on the opinion of the examining expert ophthalmologist at a single time point. Ideally, serial assessments would identify the ocular findings that are truly reversible manifestations of disease and reflective of disease activity. In addition, the majority of patients included in this study had severe global cGvHD, were on high intensity immunosuppression, and a large percentage of patients had low tear production (measured by Schirmer's tear test), as well as conjunctival and corneal surface abnormalities (measured by Oxford staining). These are, however, cGvHD patients who carry the most burden of the disease and pose major challenges in the clinic, and are, therefore, a specific focus of this research. These predictive models require validation in independent cohorts including patients with newly diagnosed or mild cGvHD for them to be applicable to other patient populations. Ideally, a prospectively designed, longitudinal multicenter study would be needed to verify these findings, in addition to having confirmation of ocular cGvHD by a comprehensive ophthalmology examination. Despite these limitations, this study was performed on a well-annotated, large sample of cGvHD patients allowing statistically meaningful analyses.

In conclusion, in a large cohort of patients with moderate to severe cGvHD, this study identified the NIH eye score and Schirmer's tear test as strongly predictive of ocular cGvHD diagnosis. A single patient self-reported item assessing dry eye symptom bother (Lee dry eye item score >2) is specific for active ocular cGvHD and should be further evaluated as a potential clinical trial outcome measure. This study also provides compelling information about the specific components of the expert ophthalmologist examination which are most associated with determining ocular cGvHD diagnosis and activity. These findings should be validated in other patient populations and will collectively help to streamline the process of ocular cGvHD diagnosis and referring post-transplant patients for evaluation by an expert ophthalmologist.

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. Jacobs R, Tran U, Chen H, et al. Prevalence and risk factors associated with development of ocular GVHD defined by NIH consensus criteria. Bone Marrow Transplant. 2012;47(11):1470-1473.
2. Dietrich-Ntoukas T, Cursiefen C, Westekemper H, et al. Diagnosis and treatment of ocular chronic graft-versus-host disease: report from the German-Austrian-Swiss Consensus Conference on Clinical Practice in chronic GVHD. Cornea. 2012;31(3):299-310.
3. Vogelsang GB, Pavletic SZ. Chronic Graft Versus Host Disease: Interdisciplinary Management: Cambridge University Press, 2009. p. 199-206.
4. Franklin RM, Kenyon KR, Tutschka PJ, Saral R, Green WR, Santos GW. Ocular manifestations of graft-vs-host disease. Ophthalmology. 1983;90(1):4-13.
5. Jabs DA, Hirst LW, Green WR, Tutschka PJ, Santos GW, Beschorner WE. The eye in bone marrow transplantation. II. Histopathology. Arch Ophthalmol. 1983;101(4):585-590.
6. Ogawa Y, Yamazaki K, Kuwana M, et al. A significant role of stromal fibroblasts in rapidly progressive dry eye in patients with chronic GVHD. Invest Ophthalmol Vis Sci. 2001;42(1):111-119.
7. Ogawa Y, Kuwana M, Yamazaki K, et al. Periductal area as the primary site for T-cell activation in lacrimal gland chronic graft-versus-host disease. Invest Ophthalmol Vis Sci. 2003;44(5):1888-1896.
8. Ogawa Y, Razzaque MS, Kameyama K, et al. Role of heat shock protein 47, a collagen-binding chaperone, in lacrimal gland pathology in patients with cGVHD. Invest Ophthalmol Vis Sci. 2007;48(3):1079-1086.
9. Bhan AK, Fujikawa LS, Foster CS. T-cell subsets and Langerhans cells in normal and diseased conjunctiva. Am J Ophthalmol. 1982;94(2):205-212.
10. Kang MH, Kim MK, Lee HJ, Lee HI, Wee WR, Lee JH. Interleukin-17 in various ocular surface inflammatory diseases. J Korean Med Sci. 2011;26(7):938-944.
11. Aronni S, Cortes M, Sacchetti M, et al. Upregulation of ICAM-1 expression in the conjunctiva of patients with chronic graft-versus-host disease. Eur J Ophthalmol. 2006;16(1):17-23.
12. He C, Lai P, Weng J, et al. Toll-like receptor 2-mediated NF-kappaB inflammatory responses in dry eye associated with cGVHD. Mol Vis. 2011;17:2605-2611.
13. Filipovich AH, Weisdorf D, Pavletic S, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. Biol Blood Marrow Transplant. 2005;11(12):945-956.
14. Jagasia MH, Greinix HT, Arora M, 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. Biol Blood Marrow Transplant. 2015;21(3):389-401.
15. Ogawa Y, Kim SK, Dana R, et al. International Chronic Ocular Graft-vs-Host-Disease (GVHD) Consensus Group: proposed diagnostic criteria for chronic GVHD (Part I). Sci Rep. 2013;3:3419.
16. Inamoto Y, Chai X, Kurland BF, et al. Validation of measurement scales in ocular graft-versus-host disease. Ophthalmology. 2012;119(3):487-493.
17. Westeneng AC, Hettinga Y, Lokhorst H, Verdonck L, van Dorp S, Rothova A. Ocular graft-versus-host disease after allogeneic stem cell transplantation. Cornea. 2010;29(7):758-763.
18. Baird K, Steinberg SM, Grkovic L, et al. National Institutes of Health chronic graft-versus-host disease staging in severely affected patients: organ and global scoring correlate with established indicators of disease severity and prognosis. Biol Blood Marrow Transplant. 2013;19(4):632-639.
19. Holladay JT. Proper method for calculating average visual acuity. J Refract Surg. 1997; 13(4):388-391.
20. Methodologies to diagnose and monitor dry eye disease: report of the Diagnostic Methodology Subcommittee of the International Dry Eye WorkShop (2007). Ocul Surf. 2007;5(2):108-152.
21. Pavletic SZ, Martin P, Lee SJ, et al. Measuring therapeutic response in chronic graft-versus-host disease: National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: IV. Response Criteria Working Group report. Biol Blood Marrow Transplant. 2006;12(3):252-266.
22. Agresti A. Categorical data analysis. New York: Wiley, 1990. p. 79-129.
23. Mehta CR, Patel NR. A Network Algorithm for Performing Fisher's Exact Test in r × c Contingency Tables. J Am Stat Assoc. 1983;78(382):427-434.
24. Lee S, Cook EF, Soiffer R, Antin JH. Development and validation of a scale to measure symptoms of chronic graft-versus-host disease. Biol Blood Marrow Transplant. 2002;8(8):444-452.
25. Lee SJ, Wolff D, Kitko C, et al. Measuring Therapeutic Response In Chronic Graft-Versus-Host Disease. Biol Blood Marrow Transplant. 2015;21(6):984-999.
26. Inamoto Y, Storer BE, Petersdorf EW, et al. Incidence, risk factors, and outcomes of sclerosis in patients with chronic graft-versus-host disease. Blood. 2013;121(25):5098-5103.
27. Imanguli MM, Atkinson JC, Mitchell SA, et al. Salivary gland involvement in chronic graft-versus-host disease: prevalence, clinical significance, and recommendations for evaluation. Biol Blood Marrow Transplant. 2010;16(10):1362-1369.
28. Jabs DA, Wingard J, Green WR, Farmer ER, Vogelsang G, Saral R. The eye in bone marrow transplantation. III. Conjunctival graft-vs-host disease. Arch Ophthalmol. 1989; 107(9):1343-1348.
29. Mitchell SA, Leidy NK, Mooney KH, et al. Determinants of functional performance in long-term survivors of allogeneic hematopoietic stem cell transplantation with chronic graft-versus-host disease (cGVHD). Bone Marrow Transplant. 2010; 45(4):762-769.