Grading criteria for chronic ocular graft-versus-host disease: Comparing the NIH eye score, Japanese dry eye score, and DEWS 2007 score

Yukako Tatematsu[1], Yoko Ogawa[1], Takayuki Abe[3], Mizuka Kamoi[1], Miki Uchino[1], Yumiko Saijo-Ban[1], Saori Yaguchi[1], Shin Mukai[1], Takehiko Mori[2], Shinichiro Okamoto[2] & Kazuo Tsubota[1]

[1]Department of Ophthalmology, Keio University School of Medicine, [2]Division of Hematology, Department of Medicine, Keio University School of Medicine, [3]Department of Preventive Medicine and Public Health, Center for Clinical Research, Keio University School of Medicine, Tokyo, Japan.

Ocular graft-versus-host disease (GVHD) is a common complication after hematopoietic stem cell transplantation (HSCT). Here we compared the diagnostic rates of ocular GVHD, including its severity, prognosis and the agreement, obtained using three grading scales: the National Institutes of Health (NIH) eye score, Japanese dry eye score, and dry eye workshop score, by retrospectively reviewing the records of 82 patients who underwent HSCT. Tear dynamics and ocular surface assessments made 6–9 months after HSCT were used to determine ocular GVHD severity with the three scales. By the three scales, ocular GVHD was diagnosed in 56 patients (68.3%), 51 patients (62.2%), and 52 patients (63.4%), respectively. The Kappa coefficient was calculated to determine the agreement between scales for diagnosing ocular GVHD. The severity progression within two years after onset was also assessed by tear dynamics and ocular surface examination. The patients were categorized as no change, improved, or progressive. The three grading scales showed good agreement (Kappa = 0.87 to 0.97) in diagnosing patients with ocular GVHD, and the scores by all three were significantly associated with the patients' prognosis (p < 0.01). We recommend that multi-center research is needed for further validation and investigation.

C hronic graft-versus-host disease (cGVHD) is a major complication of allogeneic hematopoietic stem cell transplantation (HSCT)[1], with features similar to autoimmune disorders. Symptoms usually emerge within 12 months after HSCT and may be restricted to a single organ or tissue, or they may appear at multiple sites[2]. Over the past 3 decades, the number of allogeneic hematopoietic stem cell transplantation (HSCT) procedures, and hence the cases of cGVHD, have increased.

cGVHD often affects the eyes. In this paper, we define ocular GVHD as the dry eye arising after HSCT in cGVHD patients. As humans obtain 70–80% of their sensory information from vision, improving the diagnosis[3–6] and treatment of ocular GVHD[4,7–12] will substantially enhance patients' quality of life[13,14]. Manifestations of ocular GVHD are reported in nearly 30–50% of allogeneic HSCT recipients[14–20]. Ocular GVHD typically occurs within 6 months after HSCT. The dry eye can progress rapidly to a severe state[16,17], with an increased risk for corneal ulcer and occasional perforation[21,22].

GVHD occurring more than 100 days after transplantation was previously defined as cGVHD; however, this definition has been revised[6,23]. A research group of the National Institutes of Health (NIH) suggests that a diagnosis of cGVHD requires a distinction from acute (aGVHD) and the presence of at least one diagnostic manifestation of cGVHD or at least one distinctive manifestation; the diagnosis is confirmed by pertinent biopsy or other relevant tests in the same or another organ[6]. Importantly, these criteria mean that if ocular GVHD is the only cGVHD-associated symptom, cGVHD cannot be diagnosed.

There are several diagnostic criteria for assessing the severity of ocular GVHD. The NIH eye score is a clinical scoring system proposed as NIH consensus criteria in 2005, as part of a global assessment of cGVHD severity based on the number of organs involved and the degree of impairment of the affected organs[6,23]. The Japanese dry eye score, revised in 2006, is used in Japan for ocular GVHD as well as dry eye caused by other diseases[24]. It has three parts, which assess dry eye symptomatology, tear film abnormality, and conjunctival and corneal epithelial

Table 1	Demographic and clinical characteristics of ocular GVHD patients (total n = 82)	
Ocular GVHD patients	n (%)
Gender, male/female	44 (53.7)/38 (46.3)
Median age (range)	45.5 (19–61)
Underlying disease	
Acute myeloid leukemia	23 (28.1)
Acute lymphoblastic leukemia	8 (9.8)
Myelodisplastic syndrome	20 (24.4)
Chronic myelogenous leukemia	11 (13.4)
Non-Hodgkin lymphoma	12 (14.6)
Multiple myeloma	3 (3.7)
Aplastic anemia	2 (2.4)
Others	3 (3.7)
Donor median age (range)	35.0 (17–80)
Donor gender, male/female	57 (69.5)/25(30.5)
Donor sex mismatch	
Female-male	8 (9.76)
Male-female	21 (25.6)
Donor-recipient relationship	
Related	17 (20.7)
Unrelated	65 (79.3)
Preceding aGVHD	
aGVHD (+)	59 (72.0)
aGVHD (−)	23 (28.0)
Global assessment of aGVHD severity	
Grade 0	23 (28.0)
Grade I	22 (26.8)
Grade II	29 (35.4)
Grade III	5 (6.1)
Grade IV	3 (3.7)
Stem cell source	
Bone marrow	64 (78.0)
Peripheral blood	18 (22.0)
Number of organs involved in cGVHD besides eyes*	
One organ	28 (34.1)
Two organs	14 (17.1)
Three organs	10 (12.2)
*Systemic cGVHD diagnosed by NIH consensus criteria	
Systemic cGVHD (+) ocular GVHD (+)	57 (69.5)
Systemic cGVHD (−) ocular GVHD (+)	10 (12.2)
Systemic cGVHD (−) ocular GVHD (−)	10 (12.2)
Unknown	5 (6.1)

cGVHD: chronic graft-versus-host disease, aGVHD: acute graft-versus-host disease. NIH; National Institutes of Health.
*Patients who presented cGVHD with a single or multiple organs besides the eyes were defined as having systemic cGVHD.

damage[24]. More than 4,000 HSCT procedures are performed annually in Japan, and the number is increasing. Thus, the Japanese dry eye score is widely applied. The dry eye workshop score (DEWS), reported in 2007, diagnoses dry eye based on dry eye symptomatology, tear film abnormality, conjunctival and corneal epithelial damage, and lid/meibomian gland dysfunction[7].

The NIH consensus criteria for organs beside the eyes[25,26] and the NIH eye score were previously validated[27–29]. Inamoto et al. reported that the NIH eye score shows the greatest sensitivity to symptom changes among various common scales, including the global rating of eye symptoms, Lee eye subscale, Ocular Surface Disease Index (OSDI), and Schirmer test[28]. The results obtained using all the assessment tools except for Schirmer test are correlated with both provider-reported and patient-reported changes in ocular GVHD activity. There are studies investigating the classification of ocular GVHD assessed by DEWS 2007 score[19,30–32]. However, the NIH eye score has not been compared with the Japanese dry eye score or the DEWS 2007 score with respect to the severity of dry eye diagnosed. The novel aspects of the present study are (1) that we compared the diagnostic rate and severity of ocular GVHD assessed by three grading scales: the NIH eye score, the Japanese dry eye score, and the DEWS 2007 score, and (2) that all of the ophthalmological examinations were conducted using standardized methods and schedules through the Keio BMT program; in addition, all of the HSCT patients were examined at the Keio dry eye clinic before receiving HSCT.

The purposes of this study were (1) to examine the diagnostic rates of ocular GVHD, including its severity and prognosis using the three above-mentioned grading scales at Keio University School of Medicine, and (2) to assess the agreement among the three grading scales.

To the best of our knowledge, this is the first study to compare the severity of ocular GVHD patients assessed by three grading scales, and to determine the agreement among those scales.

Results

The demographic and clinical characteristics of the 82 patients who underwent HSCT are summarized in Table 1. According to the NIH consensus criteria, 57 patients had both systemic and ocular GVHD, 10 had only ocular GVHD, 10 had neither systemic nor ocular GVHD, and the diagnosis for 5 patients was unknown.

The diagnosis and severity of ocular GVHD in each of the 82 patients who underwent HSCT was graded using three different grading scales: the NIH eye score, the Japanese dry eye score, and the DEWS 2007 score (Table 2). By these scales, ocular GVHD was diagnosed in 56 patients (68.3%), 51 patients (62.2%), and 52 patients (63.4%), respectively. The proportion of patients diagnosed with ocular GVHD was analyzed as shown in Table 3. The Kappa coefficient (K) for the proportion of patients diagnosed with ocular GVHD by the NIH eye score and Japanese dry eye score was K = 0.85 (95% CI: 0.75 to 0.98), by the NIH eye score and DEWS 2007 score was K = 0.89 (95% CI: 0.92 to 1.00), and by the Japanese dry eye score and DEWS 2007 score was K = 0.95 (95% CI: 0.79 to 0.99). Thus, all of the Kappa coefficients for the proportion of patients diagnosed with ocular GVHD showed good agreement, especially between the Japanese dry eye score and DEWS 2007 score.

A Cochran Armitage trend test was conducted to analyze the proportion of progressive cases obtained with the three measurement scales. There was a significant relationship between the scores and prognosis of ocular GVHD using all three grading scales (p < 0.0001) (Table 4).

Table 2 | Number of patients with ocular GVHD of different severities, graded according to the NIH eye score, Japanese dry eye score, and DEWS 2007 score (total n = 82)*

Grading scale (score range)	NIH eye score (0–3)		Japanese dry eye score (0–2)		DEWS 2007 score (0–4)	
	Score 0	26 (31.7%)	Score 0	31 (37.8%)	Score 0	30 (36.6%)
	Score 1	9 (11.0%)	Score 1	32 (39.0%)	Score 1	11 (14.6%)
	Score 2	42 (51.2%)	Score 2	19 (23.2%)	Score 2	24 (28.1%)
	Score 3	5 (6.1%)			Score 3	9 (11.0%)
					Score 4	8 (9.8%)

*We conducted Schirmer test with nasal stimulation to evaluate reduced reflex tearing. We defined severe dry eye as reduced reflex tearing (Schirmer test with nasal stimulation < 10 mm) or Japanese dry eye score 2, or DEWS 2007 scores 3 and 4. GVHD; graft-versus-host disease, NIH; National Institutes of Health, DEWS; dry eye work shop.

Table 3 | Proportion of patients diagnosed with ocular GVHD

Grading scale	n (total 82)	proportion	95%CI
NIH eye score			
ocular GVHD (+)	56	68.30%	57.1 to 78.1
ocular GVHD (−)	26	31.70%	
Japanese dry eye score			
ocular GVHD (+)	51	62.20%	50.8 to 72.7
ocular GVHD (−)	31	37.80%	
DEWS 2007 score			
ocular GVHD (+)	52	63.40%	52.0 to 73.8
ocular GVHD (−)	30	36.60%	

Discussion

In this study, we found that three grading scales, the NIH eye score, the Japanese dry eye score, and the DEWS 2007 score, showed statistically good agreement in the diagnostic rate and the severity of ocular GVHD. The clinical manifestations and pathology of ocular GVHD have been previously studied in Japan. The natural course[17], the baseline profiles of the ocular surface, and the tear dynamics of ocular GVHD were reported[32]. Examination of the histopathologic features of ocular GVHD revealed stromal fibroblasts in the lacrimal glands, an increased expression of HSP47[33], and morphological alteration of the conjunctival mucosal microvilli[34]. In contrast, the present study was conducted specifically to examine the severity and the agreement among different grading scales for assessing ocular GVHD.

In the present study, the NIH eye score, the Japanese dry eye score, and the DEWS 2007 score indicated ocular GVHD in 68.3%, 62.2%, and 63.4% of the HSCT patients, respectively. Similarly, Balaram et al[35] reported that 62% of HSCT patients (21 eyes out of 34 eyes) showed Schirmer test scores < 5 mm. A recent cross-sectional study of 40 allo-HSCT patients investigated the severity of ocular GVHD[19]. They reported that ocular GVHD was seen in 24 eyes (30% of all the patients), with dry eye severity of score 1 (10.0%), score 2 (2.5%), and score 3 (17.5%) diagnosed by DEWS 2007 score. Ocular GVHD was seen 63.4% of HSCT patients diagnosed by DEWS 2007 score in our study. The proportion of each score diagnosed by DEWS 2007 score in our study was score 1 (14.6%), score 2 (28.1%), score 3 (11.0%), and score 4 (9.8%) showing more severe cases. It is possible that a different frequency of ocular GVHD patients would be observed in a study group even among tertiary hospitals and from another

secondary hospital; for example, the patient profile might be different, or more patients may have a milder form of dry eye, because Keio University School of Medicine would tend to have more severe cases. Ocular GVHD accompanies cGVHD manifestations in other organs in a large percentage of cases; however, it can also be the initial and/or the only manifestation of cGVHD[30,36]. In our study, there were 10 non-systemic cGVHD patients (12.2% of all patients) who had ocular GVHD. Of these patients, some had severe manifestations of typical ocular GVHD, such as fluorescein and rose bengal staining scores of 6 points, and a Schirmer test of 0 mm. Interestingly, one of these 10 patients later developed systemic cGVHD in the skin. As our study is retrospective, it is possible that some other patients who developed cGVHD initially in the eyes and then developed systemic cGVHD immediately afterwards. However, before their next ophthalmic examination, they were included in the group of systemic cGVHD patients with ocular GVHD. Therefore, prospective research is needed to clarify the development of this disease.

It should be noted that the NIH consensus criteria, which require the involvement of an additional organ system, did not diagnose these 10 cases as cGVHD. Ocular GVHD can, however, alert clinicians to the possible presence of cGVHD in other organs[36]. Importantly, when ocular GVHD is present, but cGVHD manifestations do not occur in other organs, the NIH consensus criteria could result in a failure to diagnose patients with cGVHD, which could delay appropriate treatment.

There were some other interesting discrepancies among the scores obtained using the three grading scales. In our study, 42 patients (51.2% of all patients) were diagnosed with moderate ocular GVHD (score 2) by the NIH eye score. In other words, 75.0% of

Table 4 | Proportion of progressive cases using three grading scales

NIH eye score	x/n	Proportion (%)	95%CI	P-value#
Score 0	0/26	0	(0.0 to 13.2)	<0.0001
Score 1	1/9	11.1	(0.3 to 48.2)	
Score 2	11/42	26.2	(13.9 to 42.0)	
Score 3	4/5	80.0	(28.4 to 99.5)	
Japanese dry eye score	**x/n**	**Proportion (%)**	**95%CI**	**P-value#**
Score 0	1/31	3.2	(0.1 to 16.7)	<0.0001
Score 1	6/34	17.65	(6.8 to 34.5)	
Score 2	9/17	52.9	(27.8 to 77.0)	
DEWS 2007 score	**x/n**	**Proportion (%)**	**95%CI**	**P-value#**
Score 0	0/30	0	(0.0 to 11.6)	<0.0001
Score 1	4/12	33.3	(9.9 to 65.1)	
Score 2	2/23	8.7	(1.1 to 28.0)	
Score 3	4/9	44.4	(13.7 to 78.8)	
Score 4	6/8	75.0	(34.9 to 96.8)	

x: number of patients with progression within each score category.
n: number of patients who were diagnosed with each score.
#: p-value for Cochran Armitage trend test (two-sided).

the 56 patients diagnosed with ocular GVHD by the NIH eye score were deemed to have moderate ocular GVHD (score 2). This percentage of moderate ocular GVHD patients is similar to the rate reported by Arai et al[27]. On the other hand, we found that the ocular GVHD patients were more evenly divided into different groups of severity when assessed by the Japanese dry eye score and DEWS 2007 score than by the NIH eye score. In the NIH eye scoring system, the severity of ocular GVHD is assessed by the frequency of eye drop usage. For example, patients who need to use eye drops > 3 times per day or punctual plugs are diagnosed with moderate ocular GVHD (score 2). However, ophthalmologists instruct most of their ocular GVHD patients to use eye drops > 3 times per day. Hence, patients who have mild (score 1) or severe (score 3) ocular GVHD and use eye drops > 3 times per day could be included in the moderate ocular GVHD (score 2) group using the NIH eye score.

Some of the 42 cases diagnosed as moderate ocular GVHD by the NIH eye score were diagnosed as mild or severe cases by the Japanese dry eye score or the DEWS 2007 score. Notably, cases diagnosed as moderate or severe ocular GVHD by the NIH eye score are treated similarly[23]. Therefore, a case diagnosed as mild ocular GVHD by the other two methods but as moderate by the NIH eye score might be treated like a severe case. This could result in overtreatment, increasing the risk of cataract, glaucoma, corneal ulcer, or infection, induced by corticosteroid eye drop use. On the other hand, of the cases diagnosed as moderate ocular GVHD by the NIH eye score, 12 (14.6% of all the patients) were diagnosed as severe by the Japanese dry eye score. Since moderate and severe ocular GVHD by the NIH consensus criteria are treated similarly, the treatment in these cases would not necessarily differ as a result of their classification. However, an accurate evaluation of the severity of ocular GVHD is still important for transplantation teams, including hematologists, and ophthalmologists[29,37], since it affects the patient's health and prognosis. Thus, there were some interesting discrepancies in the assessment of ocular GVHD severity using the three grading scales, which indicate that inadequate assessments could lead to inappropriate treatment.

Jacob et al. reported the false positive and false negative rates of diagnosing ocular GVHD by Schirmer test without nasal stimulation to be 19.4% and 36.4% respectively[17,38]. In addition, the presence of dry eye may not always be due to ocular GVHD; other causes of ocular surface damage include infectious keratitis induced by the use of immunosuppression, the use of anti-glaucoma eye drops to treat corticosteroid-induced ocular hypertension, side effects of medications used to treat organ systems beside the eyes, or conditioning treatments including total body irradiation[5]. Therefore, a comprehensive ocular evaluation is recommended rather than screening with the Schirmer test to establish a diagnosis of ocular GVHD. It is important to include other ocular evaluations in the NIH eye score.

Meibomian gland dysfunction (MGD) is one of the most frequent complications of ocular cGVHD[17,32], and it is recognized more frequently in HSCT patients with cGVHD (63.0%) than in those without it (23.5%). Atrophic meibomian glands and excessive fibrosis are observed in cGVHD patients with severe dry eye, by in vivo laser confocal microscopy[39]. The extent and severity of MGD are worse in patients with immune disorders like Sjögren's syndrome and cGVHD. Especially in cGVHD, inflammatory cell infiltration and excessive fibrosis are observed around the meibomian glands, similar to the cGVHD lacrimal gland. Therefore, we should further pay much attention to the frequency, severity and relationship between dry eye and the meibomian glands.

In this study, we found that the diagnostic criteria in each of the three grading scales were similarly useful for obtaining a diagnosis of definitive ocular GVHD. However, our findings were insufficient to define a standard evaluation system for ocular GVHD that can be used by both ophthalmologists and internists. We recommend that a prospective study be performed before and after HSCT, if possible. A common system for evaluating ocular GVHD worldwide should be easy to perform, reliable, and familiar to both ophthalmologists and internists. Therefore, we recommend using the parameters proposed by ophthalmologists at the chronic ocular GVHD consensus meeting, which include the OSDI, corneal fluorescein staining score, Schirmer test value, and conjunctival injection, in future studies[40]. In this context, it is worth noting that Alvis reported that the best combination of tests to diagnose dry eye with high sensitivity, specificity, and accuracy is the OSDI, tear film break-up time, and Schirmer test[31].

Multi-center validation research by ophthalmologists is needed. At present, the chronic ocular GVHD consensus meeting of ophthalmologists to improve the sensitivity and specificity of diagnosing and treating this serious disease is in progress[40].

Methods

Patients and methods. Informed consent was obtained from all the participants, and an institutional ethics review board approval was obtained at Keio University School of Medicine (#2012-541). This study followed the guidelines of the tenets of the Declaration of Helsinki. Ethical guidelines for clinical study from the Japanese Ministry of Health, Labor, and Welfare indicate the studies which do not involve biological tissue and which involve reviewing medical records retrospectively, researchers do not need to obtain written informed consent from patients. Following the guidelines of the ethics committees, we posted a detailed written guideline and ethical statement of the present study in our outpatient clinic of ophthalmology. The notice included the following factors: background of the study, purpose of the study, study design, privacy policy, freedom to withdraw, inclusion and exclusion criteria, the factors assessed in the medical records, advantage and disadvantage of participating the study, disclosure of the data, presenting the data at a conference or in a journal, and contact information.

Inclusion criteria for this study were patients who received bone marrow transplantation and peripheral blood stem cell transplantation at the Division of Hematology, Keio University School of Medicine, Tokyo, Japan between April 2004 and January 2010. Patients 20 years or older who underwent HSCT for the first time and survived 100 days after the transplant, and those with sustained donor engraftment were included. The severity of ocular GVHD is affected by multiple factors, including donor-recipient gender difference and the stem cell source[41,42]. In this study, we included ocular GVHD and excluded any patients with other risk factors for dry eye, so the severity of ocular GVHD could be compared under standardized conditions. In particular, patients with a history of Sjögren's syndrome, rheumatic disease, diabetes mellitus, Graves' disease, other systemic or ocular diseases including glaucoma, a history of ocular surgery including LASIK, contact lens use, or drug use including psychotropic drugs, were excluded. Patients who had dry eye before HSCT were also excluded. Patients who underwent cord blood stem cell transplantation were excluded. In total, the records of 82 HSCT patients (38 women and 44 men) with a median age of 45.5 years (range, 19–61) were reviewed retrospectively.

All of the HSCT patients underwent standardized clinical and ophthalmological evaluations at the dry eye clinic at Keio University School of Medicine before HSCT and 3, 6, 9, 12, 18, and 24 months after HSCT, as arranged by the Keio BMT program transplant internist. The Keio BMT program was begun in 1994 to establish a collaboration between internists and Ophthalmologists[17]. We used the data obtained approximately 6–9 months after HSCT to assess the severity of the ocular GVHD. We also assessed the degree of ocular GVHD 3–6 months after the onset of dry eye, for up to 24 months, to categorize the patients as no change, improved, or progressive. We previously found that severe dry eye with diminished reflex tearing appeared 36 months after the onset of dry eye[17]. A progressive case was defined as the worsening of objective ocular signs and/or the requirement for additional therapy. Some progressive cases were evaluated every week as additional examinations.

Tear function examinations and ocular surface vital staining. To compare patients with three measurement scales, we used the values for tear function and ocular surface vital staining abnormality based on the Japanese diagnostic criteria, as reported previously[24,43]. For ocular surface double staining, 2 μl of a preservative-free mixture of 1% rose bengal and 1% fluorescein was instilled into the conjunctival sac with a micropipette, as reported previously[44]. To determine the tear break up time, the interval between the last complete blink and the appearance of the first corneal black spot in the stained tear film was measured three times, and the mean value was calculated. The Schirmer 1-test was performed without topical anesthesia using standardized strips of filter paper (Alcon Inc., Fort Worth, TX, USA). Readings were reported in millimeters of wetting for 5 min.

Grading scales. *NIH eye score.* In the NIH consensus, the ocular criteria for diagnosis are defined as, "new ocular sicca documented by low Schirmer test values with a mean value of both eyes ≤ 5 mm at 5 minutes or a new onset of keratoconjunctivitis sicca by slit-lamp examination with mean values of 6 to 10 mm on the Schirmer test is sufficient for the diagnosis of chronic GHVD if accompanied by distinctive manifestations in at least 1 other organ." The NIH eye score has a range of 0–3 (0 =

non dry eye, 1 = mild dry eye, 2 = moderate dry eye, 3 = severe dry eye) (Appendix 1; Supplemenary information)[23].

Japanese dry eye score. The Japanese dry eye criteria for diagnosis are: (1) disturbance of the tear film (Schirmer test ≤ 5 mm or tear film breakup time ≤ 5 seconds); (2) conjunctivocorneal epithelial damage (fluorescein staining score ≥ 3 points or rose bengal staining score ≥ 3 points); and (3) dry eye symptomatology. The presence of all three criteria is necessary for a diagnosis of definite dry eye disease (Appendix 2; Supplementary information)[24]. The severity of the dry eye with this system is scored with a range of 0–2 (0 = non dry eye, 1 = mild dry eye, 2 = severe dry eye). A score of 0 indicates non dry eye presenting no manifestations/symptoms, a score of 1 indicates symptoms, Schirmer test ≤ 5 mm, fluorescein score < 3 points, and rose bengal score < 3 points; and a score of 2 indicates symptoms, Schirmer test ≤5 mm, fluorescein score ≥ 3 points, and rose bengal score ≥ 3 points[17,45].

DEWS 2007 score. Third, the DEWS 2007 score has a score range of 0–4 (0 = non dry eye, 1 = mild dry eye, 2 = moderate dry eye, 3 = severe dry eye, 4 = very severe dry eye), and is determined from 9 parameters, including symptoms, Schirmer test score, tear film breakup time, and abnormalities in the conjunctiva, cornea, tear, lid, and meibomian glands (Appendix 3; Supplementary information)[7].

Since each grading scale has a different score range, we evaluated the reduction in reflex tearing by conducting Schirmer tests with nasal stimulation. We defined severe dry eye as reduced reflex tearing (Schirmer test with nasal stimulation < 10 mm), as Japanese dry eye score 2, or as DEWS 2007 scores 3 and 4. Therefore, there was overlap in the patients with Schirmer test with nasal stimulation ≥ 10 mm, between those with NIH score 2 and Japanese dry eye score 1; and in the patients with Schirmer test with nasal stimulation < 10 mm, between the patients with NIH score 2 and Japanese dry eye score 2.

Statistical analyses. Descriptive statistics, including the median and range for continuous variables and the percentage and frequencies of categorical variables, were calculated in assessing the demographic and clinical characteristics of ocular GVHD patients. Agreement of binary diagnoses between pairs of the three grading scales (NIH eye score, Japanese dry eye score, and DEWS 2007 score) was evaluated with the Kappa coefficient, and its 95% confidence intervals (95% CI) were estimated.

The linear trend in an event proportion across a factor was tested by the exact Cochran-Armitage trend test. Significance levels for all tests were two-sided and 0.05. All data were analyzed with SPSS version 19.0 (SPSS Inc., Chicago, IL, USA) and the SAS version 9.3 (Cary, NC, USA). A p-value < 0.05 was considered statistically significant.

1. Lee, S. J. *et al.* Severity of chronic graft-versus-host disease: association with treatment-related mortality and relapse. *Blood.* **100**, 406–414 (2002).
2. Ferrara, J. L., Levine, J. E., Reddy, P. & Holler, E. Graft-versus-host disease. *Lancet* **373**, 1550–1561 (2009).
3. Espana, E. M., Shah, S., Santhiago, M. R. & Singh, A. D. Graft versus host disease: clinical evaluation, diagnosis and management. *Graefes Arch Clin Exp Ophthalmol* **251**, 1257–1266 (2013).
4. Hessen, M. & Akpek, E. K. Ocular graft-versus-host disease. *Curr Opin Allergy Clin Immunol* **12**, 540–547 (2012).
5. Dietrich-Ntoukas, T. *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* **31**, 299–310 (2012).
6. Filipovich, A. H. Diagnosis and manifestations of chronic graft-versus-host disease. *Best Pract Res Clin Haematol* **21**, 251–257 (2008).
7. The definition and classification of dry eye disease: report of the Definition and Classification Subcommittee of the International Dry Eye WorkShop (2007). *Ocul Surf* **5**, 75–92 (2007).
8. Cutler, C. *et al.* Rituximab for steroid-refractory chronic graft-versus-host disease. *Blood* **108**, 756–762 (2006).
9. Ogawa, Y. *et al.* Successful treatment of dry eye in two patients with chronic graft-versus-host disease with systemic administration of FK506 and corticosteroids. *Cornea* **20**, 430–434 (2001).
10. Ogawa, Y. *et al.* Autologous serum eye drops for the treatment of severe dry eye in patients with chronic graft-versus-host disease. *Bone Marrow Transplant* **31**, 579–583 (2003).
11. Wang, Y. *et al.* Ocular surface and tear functions after topical cyclosporine treatment in dry eye patients with chronic graft-versus-host disease. *Bone Marrow Transplant* **41**, 293–302 (2008).
12. Yaguchi, S. *et al.* Surgical management of lacrimal punctal cauterization in chronic GVHD-related dry eye with recurrent punctal plug extrusion. *Bone Marrow Transplant* **47**, 1465–1469 (2012).
13. Fraser, C. J. *et al.* Impact of chronic graft-versus-host disease on the health status of hematopoietic cell transplantation survivors: a report from the Bone Marrow Transplant Survivor Study. *Blood* **108**, 2867–2873 (2006).
14. Pidala, J., Anasetti, C. & Jim, H. Quality of life after allogeneic hematopoietic cell transplantation. *Blood* **114**, 7–19 (2009).
15. Calissendorff, B., el Azazi, M. & Lonnqvist, B. Dry eye syndrome in long-term follow-up of bone marrow transplanted patients. *Bone Marrow Transplant* **4**, 675–678 (1989).
16. Mencucci, R. *et al.* Ophthalmological aspects in allogenic bone marrow transplantation: Sjogren-like syndrome in graft-versus-host disease. *Eur J Ophthalmol* **7**, 3–18 (1997).
17. Ogawa, Y. *et al.* Dry eye after haematopoietic stem cell transplantation. *Br J Ophthalmol* **83**, 1125–1130 (1999).
18. Shikari, H., Antin, J. H. & Dana, R. Ocular graft-versus-host disease: a review. *Surv Ophthalmol* **58**, 233–251 (2013).
19. Vanathi, M. *et al.* Ocular surface evaluation in allogenic hematopoietic stem cell transplantation patients. *Eur J Ophthalmol* **24**, 655–666 (2014).
20. Riemens, A. *et al.* Impact of ocular graft-versus-host disease on visual quality of life in patients after allogeneic stem cell transplantation: questionnaire study. *Acta Ophthalmol* **92**, 82–87 (2014).
21. Inagaki, E. *et al.* Four cases of corneal perforation in patients with chronic graft-versus-host disease. *Mol Vis* **17**, 598–606 (2011).
22. Lavid, F. J., Herreras, J. M., Calonge, M., Saornil, M. A. & Aguirre, C. Calcareous corneal degeneration: report of two cases. *Cornea* **14**, 97–102 (1995).
23. Filipovich, A. H. *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* **11**, 945–956 (2005).
24. Shimazaki, J. Definition and diagnosis of dry eye 2006. *Journal of the Eye* **24**, 181–184 (2007).
25. Aisa, Y. *et al.* Validation of NIH consensus criteria for diagnosis and severity-grading of chronic graft-versus-host disease. *Int J Hematol* **97**, 263–271 (2013).
26. Jacobsohn, D. A. *et al.* Correlation between NIH composite skin score, patient-reported skin score, and outcome: results from the Chronic GVHD Consortium. *Blood* **120**, 2545–2552 (2012).
27. Arai, S. e. *et al.* Global and organ-specific chronic graft-versus-host disease severity according to the 2005 NIH Consensus Criteria. *Blood* **118**, 4242–4249 (2011).
28. Inamoto, Y. *et al.* Validation of measurement scales in ocular graft-versus-host disease. *Ophthalmology* **119**, 487–493 (2012).
29. Mitchell, S. A. *et al.* A multicenter pilot evaluation of the National Institutes of Health chronic graft-versus-host disease (cGVHD) therapeutic response measures: feasibility, interrater reliability, and minimum detectable change. *Biol Blood Marrow Transplant* **17**, 1619–1629 (2011).
30. Nassar, A., Tabbara, K. F. & Aljurf, M. Ocular manifestations of graft-versus-host disease. *Saudi J Ophthalmol* **27**, 215–222 (2013).
31. Alves, M. *et al.* Comparison of diagnostic tests in distinct well-defined conditions related to dry eye disease. *PLoS One* **9**, e97921 (2014).
32. Wang, Y. *et al.* Baseline profiles of ocular surface and tear dynamics after allogeneic hematopoietic stem cell transplantation in patients with or without chronic GVHD-related dry eye. *Bone Marrow Transplant* **45**, 1077–1083 (2010).
33. Ogawa, Y. *et al.* Role of heat shock protein 47, a collagen-binding chaperone, in lacrimal gland pathology in patients with cGVHD. *Invest Ophthalmol Vis Sci* **48**, 1079–1086 (2007).
34. Tatematsu, Y. *et al.* Mucosal microvilli in dry eye patients with chronic GVHD. *Bone Marrow Transplant* **47**, 416–425 (2012).
35. Balaram, M. & Dana, M. R. Phacoemulsification in patients after allogeneic bone marrow transplantation. *Ophthalmology* **108**, 1682–1687 (2001).
36. Kim, S. K. & Dunn, J. P. *Chronic ocular graft versus host disease.* Interdisciplinary Management (Cambridge University Press, Cambridge, UK 2009).
37. Dignan, F. L. *et al.* Organ-specific management and supportive care in chronic graft-versus-host disease. *Br J Haematol* **158**, 62–78 (2012).
38. Jacobs, R. *et al.* Prevalence and risk factors associated with development of ocular GVHD defined by NIH consensus criteria. *Bone Marrow Transplant* **47**, 1470–1473 (2012).
39. Ban, Y. *et al.* Morphologic evaluation of meibomian glands in chronic graft-versus-host disease using in vivo laser confocal microscopy. *Mol Vis* **17**, 2533–2543 (2011).
40. Ogawa, Y. *et al.* International Chronic Ocular Graft-vs-Host-Disease (GVHD) Consensus Group: proposed diagnostic criteria for chronic GVHD (Part I). *Sci Rep* **3**, 3419; doi:10.1038/srep03419 (2013).
41. Kamoi, M. *et al.* Donor-recipient gender difference affects severity of dry eye after hematopoietic stem cell transplantation. *Eye (Lond)* **25**, 860–865 (2011).
42. Uchino, M. *et al.* Comparison of stem cell sources in the severity of dry eye after allogeneic haematopoietic stem cell transplantation. *Br J Ophthalmol* **6**, 34–37 (2012).
43. Uchino, M. *et al.* Prevalence of dry eye disease and its risk factors in visual display terminal users: the Osaka study. *Am J Ophthalmol* **156**, 759–766 (2013).
44. Toda, I. & Tsubota, K. Practical double vital staining for ocular surface evaluation. *Cornea* **12**, 366–367 (1993).
45. Tsubota, K. *et al.* Three different types of dry eye syndrome. *Cornea* **13**, 202–209 (1994).

Acknowledgments

Author contributions

Y.T. and Y.O. wrote the main manuscript text, and Y.T. and Y.O. prepared tables 1–4. T.A. performed statistical analyses. All authors (Y.T., Y.O., T.A., M.K., M.U., Y.S., S.Y., S.M., T.M., S.O. and K.T.) participated in collecting and analyzing data. All authors reviewed the manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Tatematsu, Y. *et al.* Grading criteria for chronic ocular graft-versus-host disease: Comparing the NIH eye score, Japanese dry eye score, and DEWS 2007 score. *Sci. Rep.* **4**, 6680; DOI:10.1038/srep06680 (2014).