Ocular surface indicators and biomarkers in chronic ocular graft-versus-host disease: a prospective cohort study

Alexandra A. Pietraszkiewicz1 · Debbie Payne1 · Maria Abraham2 · Angel Garced1 · Krishna C. Devarasetty2 · Megan Wall2 · Supriya M. Menezes2 · Sveti Ugarte3 · Filip Pirsl4 · Sencer Goklemez4 · Frederick L. Ferris1 · John Barrett5 · Minoo Battiwalla5 · Richard W. Childs5 · Steven Z. Pavletic4 · Rachel J. Bishop1

Received: 1 August 2020 / Revised: 30 January 2021 / Accepted: 19 February 2021 / Published online: 8 March 2021
© The Author(s), under exclusive licence to Springer Nature Limited 2021. This article is published with open access, corrected publication 2021

Abstract
This longitudinal cohort study compared ocular surface indicators in forty allogeneic hematopoietic stem cell transplant (HSCT) subjects with twenty healthy controls at baseline and identified changes in ocular graft-versus-host disease (oGVHD). Outcome measures included: Ocular Surface Disease Index (OSDI), tear osmolarity, Schirmer’s test, Oxford corneal staining score, tear break-up time (TBUT), and tear and serum biomarkers (IFN-γ, IL-10, MMP-9, IL-12, IL-13, IL-17α, IL-1β, IL-2, IL-4, IL-6, IL-8, CXCL10, MCP-1, MIP-1α, RANTES, TNF-α). At baseline the HSCT group had higher median Oxford corneal staining score (1.7 vs. 0.0; \( P < 0.0001 \)), higher tear TNF-α (20.0 vs. 11.2 pg/mL; \( P < 0.0001 \)), lower tear RANTES (70.4 vs. 190.2 pg/mL; \( P < 0.0001 \)), higher serum IL-8 (10.2 vs. 4.5 pg/mL; \( P = 0.0008 \)), and higher serum TNF-α (8.7 vs. 4.2 pg/mL; \( P < 0.0001 \)). The incidence of oGVHD was 62% and associated changes included increased Oxford corneal staining score (4.6 vs. 1.8, \( P = 0.0001 \)), decreased Schirmer’s test (3.0 vs. 10.0; \( P < 0.0001 \)), and decreased TBUT (4.7 vs. 9.0 s; \( P = 0.0004 \)). Baseline differences in ocular surface indicators suggest a tendency toward ocular dryness in individuals with hematologic disorders preparing for HSCT. Individuals who developed oGVHD showed changes in corneal staining score, Schirmer’s test, and TBUT.

Introduction
Allogeneic Hematopoietic Stem Cell Transplantation (HSCT) is a potentially curative treatment for many hematologic diseases [1]. A major cause of morbidity and mortality following HSCT is graft-versus-host-disease (GVHD), a T-cell mediated immunologic response of the donor graft to host antigens [2]. The most important of these are human leukocyte antigens, found on virtually all nucleated cells [2]. Systemic chronic GVHD (cGVHD) occurs in 30–70% of transplant recipients; ocular manifestations occur in up to 90% of patients with cGVHD [1]. The main targets of disease in ocular chronic GVHD (oGVHD) are structures of the ocular surface, including the eyelids, lacrimal gland, conjunctiva, and cornea [1]. Fibrosis and destruction of ocular surface tissues result in the clinical signs and symptoms of keratoconjunctivitis sicca, blepharitis, and Meibomian gland dysfunction [1].

Baseline ophthalmological evaluation of patients undergoing HSCT has been recommended by the National Institutes of Health (NIH) Consensus Conference (2014), the First International Chronic Ocular GVHD Consensus Group (2013), and the German–Austrian–Swiss Consensus Conference (2012) [3–5]. Dry eye disease (DED), a multifactorial disorder characterized by loss of homeostasis of the tear film accompanied by ocular symptoms [6], occurs in ~50% of HSCT patients at baseline [7, 8], with alterations in ocular surface parameters developing after HSCT...
regardless of oGVHD diagnosis [9]. Studies have explored various ocular surface parameters [10] in an attempt to identify risk factors associated with the development of oGVHD and increased severity of oGVHD [11–14] and to more effectively understand the course of disease over time [15]. Tear and serum cytokines have been investigated in the hopes of establishing more reliable predictive, diagnostic, and prognostic biomarkers [16–18].

In this prospective study, 40 individuals with a hematologic disorder or malignancy who were scheduled to receive a HSCT and 20 healthy controls underwent comprehensive ophthalmologic evaluations including administration of the Ocular Surface Disease Index (OSDI) Questionnaire, a detailed ocular surface assessment, and tear and serum biomarker analysis. Baseline findings were compared between the groups. The HSCT cohort was followed over time to identify risk factors for and markers of oGVHD development. In addition, for those HCST patients who received at least one year of ophthalmic follow-up, comparison of ocular indicators and biomarkers was made between the month 12 and baseline visits.

**Subjects and methods**

**Patients and healthy controls**

The study followed the tenets of the Declaration of Helsinki and was approved by the Institutional Review Board of the National Eye Institute (NEI), NIH. Written, informed consent was obtained from each patient after explanation of the nature and possible consequences of the study. Since this was an exploratory study, no formal sample size estimation was performed. Participants were enrolled from HSCT treatment protocols at the National Cancer Institute (NCI) and the National Heart, Lung and Blood Institute (NHLBI) at NIH. Participants had to be at least 18 years of age and had to be scheduled for a HSCT under an NIH protocol. Participants had a range of primary diagnoses (Table 1). Exclusion criteria included a history of ocular problems that could interfere with the natural history of their response to treatment with HSCT (significant dry eye disease, use of cyclosporine eye drops in the past 30 days, conjunctival scarring, etc.), and known allergies to dilating or anesthetic eye drops.

The intended HSCT cohort included 40 individuals. However, four individuals did not receive a transplant and two died within one month of transplant; the remaining 34 individuals comprised the analytic HSCT cohort. Of those, 30 individuals had at least 12 months of follow-up, comprising the longitudinal HSCT cohort. Healthy controls \((n = 20)\) were recruited through the NEI consult service (at NIH) and were at least 18 years of age without significant systemic or ocular disease.

**Clinical examination and sample collection**

Participants underwent the following tests as part of their baseline and subsequent evaluations: medical and ophthalmic history, OSDI scoring, best-corrected visual acuity, tonometry, Schirmer’s test with anesthesia, tear break-up time (TBUT), a detailed anterior segment examination, conjunctival and corneal surface examinations including vital dye staining with lissamine green and fluorescein, and a dilated fundus examination. The anterior segment examination included evaluation of the lacrimal gland, eyelids including Meibomian glands, conjunctiva, cornea, anterior
chamber, and iris. Corneal and conjunctival staining was graded using the Oxford scheme, with two modifications. First, a decimalized scale allowed graders optimal flexibility in assigning values to each segment (cornea, nasal conjunctiva, and temporal conjunctiva). Second, a value of zero indicated no punctate keratopathy, which differs from the standard grading scale which allows for a single punctate defect in the grade of zero. Each segment received a score of zero to five, with a total Oxford score of zero to 15. Tear fluid was collected and analyzed for osmolarity using the Tearlab Osmolarity System (San Diego, CA) and analyzed for biomarkers by Allergan, Inc. Serum obtained from blood samples (~5cc) was also analyzed for biomarkers (Allergan, Inc.).

Examinations were performed at baseline (between 10 and 112 days before HSCT), and at months 1, 3, 6, 9, and 12 after HSCT, with an optional extension to 18 and 24 months, and then annually for up to 5 years. Participants received all standard ocular therapies to manage dry eye and oGVHD as clinically indicated, including topical lubrication, topical cyclosporine, topical steroids, punctal occlusion, eyelid compresses and scrubs, and environmental modifications such as humidifiers. Participants were allowed to receive treatment from their home ophthalmologists throughout the study. Diagnosis of Dry Eye Disease (DED) was made according to the 2007 Report of the International Dry Eye Workshop (DEWS), the current guideline at time of study initiation [19]. Diagnosis of oGVHD was made according to the NIH consensus criteria (2014) [3]. Global scoring of chronic GVHD and organ-specific scoring were performed according to the same guidelines. For patients with DED at baseline, evidence of increased staining compared to baseline with a decrease in Schirmer’s test to a value between 6 and 10 mm, or a decrease in Schirmer’s test to below 5 mm was required for diagnosis of oGVHD.

Statistical analysis

For tear and serum biomarker analyses, a Bonferroni-corrected type I error rate of 0.16% was used. For the remaining analyses, a type I error rate of 5% was used and no adjustments were made for multiplicity. The following analyses were performed:

- Comparison of baseline ocular factors and tear and serum biomarkers between the intended HSCT group (n = 40) and controls (n = 20) using a nonparametric Wilcoxon rank-sum test.
- Among those who developed oGVHD (n = 21): comparison of ocular factors and tear and serum biomarkers at time of diagnosis of oGVHD vs. baseline, using a Wilcoxon signed-rank test. This analysis did not account for varying times to diagnosis of oGVHD. If a participant was diagnosed with oGVHD in both eyes at the same visit, both eyes were included in the analysis. If they were diagnosed with oGVHD in one eye, even if the other eye was diagnosed at a later visit, only the first affected eye was included in the analysis. Unaffected eyes were not included in the analysis.

Results

Baseline characteristics

This study enrolled 40 participants who intended to receive allogeneic HSCT (intended HSCT cohort) and 20 healthy controls. Description of study participants can be found in Table 1. Compared to healthy controls, the intended HSCT cohort had higher baseline median Oxford corneal staining...
scores, lower tear RANTES, higher tear TNF-α, higher serum IL-8, and higher serum TNF-α (Table 2). The remaining ocular factors and tear and serum biomarkers were not significantly different between the two groups at baseline (Supplementary Table S1).

Four participants in the intended HSCT cohort did not receive the planned transplant and were excluded from subsequent analyses. Two HSCT recipients received no ophthalmological follow-up after transplant due to death. Of the remaining 34 participants (i.e., the analytic cohort), thirty participants received at least 12 months of follow-up; 21 received at least 2 years of follow up, and 16 received at least three years of follow-up (Supplementary Table S2).

Rates of GVHD

oGVHD was diagnosed in 21 (62%) of the analytic HSCT cohort (Table 3). Most who developed oGVHD did so within the first 12 months (19 [91%]), with a mean time to diagnosis of ~6 months. Chronic GVHD (cGVHD) was diagnosed in 24 (71%) of the analytic HSCT cohort (Supplementary Table S3). The most commonly affected initial organ was the eye (14, [58%]) followed by the mouth (7, [29%]).

The global NIH cGVHD score at the time of oGVHD diagnosis mirrored the initial ocular score in most cases, consistent with the finding that the eye was the first organ affected for most participants who developed cGVHD (Table 4A). Examination of the cross-frequencies between the maximum global scores and the maximum ocular score revealed that two of the four cases of severe cGVHD were...
driven by severe oGVHD (Table 4B). Four of the twelve cases of moderate cGVHD were driven by moderate ocular involvement.

Ocular surface factors and biomarkers in ocular GVHD

Compared with baseline values, patients who developed oGVHD had lower Schirmer’s test and increased Oxford corneal staining score at time of diagnosis; this was expected as these were defining features of the diagnosis (Table 5). In addition, TBUT was significantly decreased at time of diagnosis of oGVHD compared with baseline (Table 5). OSDI and tear osmolarity were not significantly different. None of the tear and serum biomarkers were significantly different between baseline and time of oGVHD diagnosis (Supplementary Table S4).

Receiver operating characteristic (ROC) curve analysis

Among all parameters evaluated, Oxford corneal staining score, tear osmolarity, and serum biomarkers CXCL10 and TNF-α had an AUC > 0.7 for oGVHD diagnosis (Fig. 1). The optimal cutoffs, sensitivities, specificities, Positive Predictive Values (PPV), and Negative Predictive Values (NPV) are presented in Supplementary Table S5.

Trends in the HSCT Group in the first 12 months

The HSCT cohort was examined longitudinally, comparing ocular surface factors and biomarkers at 12 months to baseline (Supplementary Table S6). Oxford corneal staining score increased and TBUT decreased. When assessed nonparametrically, median tear IL-8 was significantly increased. Trends in these variables over 12 months of follow-up are shown in Fig. 2a, b.

Risk factors for GVHD

Older age (>40 years) was found to increase the odds of developing oGVHD in the unadjusted model (OR 7.3, P = 0.01). Older age was also found to significantly increase the odds of developing cGVHD in both models (Supplementary Table S7). A myeloablative conditioning regimen appeared to moderately increase the odds of developing acute GVHD (aGVHD) compared to a non-myeloablative conditioning regimen in the unadjusted model. No other baseline characteristics significantly increased the odds of developing ocular, acute systemic, or chronic systemic GVHD.

Discussion

This study demonstrates significant differences in dry eye indicators and selected tear and serum biomarkers in patients with hematological diseases planning to undergo
Table 5 Ocular GVHD cohort: comparison of ocular surface indicators at time of diagnosis vs. baseline.

| Ocular surface parameters | N^4 | oGVHD at Baseline Median (Range) | oGVHD at Diagnosis Median (Range) | Wilcoxon p value |
|---------------------------|-----|---------------------------------|-----------------------------------|-----------------|
| OSDI                      | 10  | 2.1 (0.0, 12.5)                 | 2.5 (0.0, 44.0)                   | 0.47            |
| Oxford Corneal Staining Score | 23  | 1.8 (0.1, 5.1)                 | 4.6 (0.4, 9.8)                   | 0.0001          |
| Schirmer’s Test (mm)      | 23  | 10.0 (3.0, 25.0)               | 3.0 (0.0, 10.0)                   | <0.0001         |
| Tear Osmolarity (mOsm/L)  | 23  | 308.0 (296.0, 345.0)           | 309.0 (288.0, 343.0)             | 0.86            |
| TBUT (s)                  | 23  | 9.0 (2.7, 13.0)                | 4.7 (2.3, 9.7)                   | 0.0004          |

^4N corresponds to the number of participants for OSDI; otherwise, corresponds to number of eyes.

OSDI ocular surface disease index, TBUT tear break-up time.

allogeneic HSCT compared to healthy controls. Approximately 38% of HSCT patients demonstrated clinically significant DED at their baseline visit compared to 25% of controls, consistent with previous studies [7, 8]. The HSCT group differed significantly from healthy controls with respect to corneal staining score. Although there has been some debate surrounding how sensitive corneal staining is for DED [21], it provides a robust method for quantifying epithelial surface damage [22]. This damage may be due to the combined effects of conditioning regimens, chemotherapy, immunosuppressive therapy, the hospital environment and other factors not accounted for. We observed a trend toward increased tear osmolarity at baseline in the HSCT group. These findings emphasize the importance of baseline ophthalmologic evaluations of patients prior to HSCT. Diagnosing oGVHD and distinguishing it from DED in the months following HSCT can be challenging if baseline evaluations are unavailable [15, 23]. DED is also known to occur in the majority of patients with cGVHD, regardless of whether they meet strict criteria for oGVHD [23, 24].

An additional challenge in diagnosing oGVHD is the existence of multiple diagnostic and severity scoring criteria. In 2014 the NIH Consensus Development Project removed Schirmer’s testing from organ severity scoring due to poor correlation with symptom change over time [3]. Minimal NIH criteria for diagnosing oGVHD retained Schirmer’s testing, but no longer required distinctive manifestations of cGVHD in another organ [3]. We used the NIH criteria for this study, allowing us to compare findings with other recent studies that also used these criteria [8, 10, 25].

Sixty two percent of HSCT patients developed oGVHD, consistent with a previous NIH-based study [26]. The incidence of oGVHD in the literature ranges from 12-55% [8, 10, 25, 27]. We report a significant decrease in TBUT in those who develop oGVHD, consistent with other studies [10, 27]. Schirmer’s and corneal staining scores have been reported as sensitive markers of oGVHD in multiple studies, as in ours, but carry the same qualifier of commonly being part of the defining criteria [10, 26, 27]. Oxford corneal staining had the highest Positive Predictive Value (79%) for diagnosis of oGVHD at a cutoff value of 3.2. Some groups have shown a strong correlation between increased tear osmolarity and oGVHD [10, 28], while others have failed to show a relationship [29, 30].
there was no significant difference in mean tear osmolarities at time of oGVHD diagnosis compared with baseline in our study, ROC analysis rendered a cut-off value of 306 mOsm/L, suggesting that tear osmolarity could contribute to oGVHD diagnosis. No significant difference was observed in OSDI scores between intended HSCT patients and controls at baseline, and OSDI scores did not increase significantly from baseline to time of diagnosis of oGVHD. Of note, OSDI often poorly correlates with clinical signs of DED due to changes in ocular surface sensitivity [21]; despite these limitations, some groups have shown a positive correlation between OSDI and objective measures of oGVHD [15, 27].

Previous studies have shown that risk factors for cGVHD include prior history of aGVHD, older age, female donor to male recipient, mismatched or unrelated donors, donor lymphocyte infusion (DLI), and use of peripheral blood stem cells [31, 32]. In this study, older age was associated with increased odds of developing oGVHD and cGVHD, and a myeloablative conditioning regimen increased the odds of developing aGVHD compared to a non-myeloablative regimen. Other baseline risk factors did not show statistically significant associations, possibly due to the small sample size or differences in study populations. Interestingly, presence of abnormal baseline ocular surface parameters could not predict subsequent development of oGVHD, a finding that may have practical implications if confirmed in future studies.

DED is an inflammatory condition that involves cytokine production by the lacrimal gland [6, 33]. As a result, cytokine and matrix metalloproteinase levels in tears have been investigated in the hopes of discovering non-invasive,

**Fig. 2** Significant ocular surface parameters and biomarkers over time (Means). a Ocular surface parameters. Top: Oxford corneal staining per the Oxford grading scheme as detailed in the methods section, and bottom: Tear Break Up Time (TBUT) in seconds. b Tear biomarker Interleukin 8 (IL-8) in pg/mL.
sensitive, and predictive markers of oGVHD. Methods of tear collection, assays, and software used for analysis differ among studies, making it challenging to compare results. Ranges of reported values vary considerably, and there are few population-based studies forming established normal levels. Tear fluid MMP-9 is an established marker of ocular surface inflammation [25, 34–36]. It was found to be elevated in the oGVHD group at time of diagnosis and in the HSCT group overall at 12 months post-transplant, though with significant variability so no p-values could be calculated. This study also demonstrated an increase in tear and serum CXCL10 (previously called IP-10) in the oGVHD group at time of diagnosis and in the HSCT group overall at 12 months post-transplant; again, p-values were not calculated due to the wide range of values. Others have shown CXCL10 to predict oGVHD [17]. Tear IL-8 was significantly increased at baseline and at 12 months post-transplant but did not reach statistical significance in the oGVHD cohort and may be elevated in hematologic disorders in general.

Though most studies focus on tear biomarkers due to their proximity to the tissue of interest and ease of procurement, we also examined serum biomarkers in order to compare systemic against organ-specific responses. Serum TNF-α had the highest sensitivity (88%) and NPV (90%) for oGVHD of all the biomarkers studied. Tear and serum TNF-α were significantly higher in the intended HSCT group at baseline compared to the control group, so these also may be elevated in hematologic disorders in general and would require validation as oGVHD-specific biomarkers. TNF-α is a pro-inflammatory cytokine that promotes T cell activation [37], and has been found to correlate with ocular surface parameters and oGVHD severity in other studies [21, 38]. Serum MIP-1α and IL-10 were increased in the oGVHD group at time of diagnosis. These are additional candidates for future investigation.

Limitations of the present report include small sample size, an incomplete data set due to gaps in follow-up of the sickest patients, and an inability to collect tear samples in the most severe cases of oGVHD. These problems also reflect real-world clinical practice. Due to the small sample size, the results from the analyses performed are considered exploratory and not conclusive; the results may be used to inform hypotheses of future studies. Considerable variability and wide confidence intervals were observed in the results presented, which can be attributed to the limited sample size. Strengths of this report include prospective longitudinal study design, comparison of key ocular surface examination indicators with a pre-transplant baseline, and baseline comparisons against a control group. We examined ocular surface indicators, tear and serum biomarkers, and well-annotated clinical transplant and systemic GVHD data at predetermined time points [39]. We show that there are major alterations in the ocular surface and biomarkers prior to HSCT, underscoring the importance of obtaining a baseline eye examination in these patients. This study highlights the importance of TBUT as a clinical marker of oGVHD, in addition to the defining criteria of Oxford corneal staining and Schirmer’s testing. These clinical findings contribute to a large body of research whose limitations include differing defining criteria for oGVHD across studies. Biomarkers hold potential to improve the accuracy of diagnosing oGVHD and better tracking response to treatment. In our study, no specific biomarkers were associated with the development of oGVHD at a statistically significant level. However, certain biomarkers showed possible correlation with oGVHD including tear and serum MMP-9 and CXCL10; tear IL-8; and serum TNF-α, IL-10, and MIP-1α. Each of these are known markers in the GVHD inflammatory pathway and have been linked to DED or to oGVHD in prior studies. Together with ocular surface indicators, these biomarkers represent potential targets for future clinical research.

Acknowledgements We thank Fran Hakim, Ph.D., for sharing expertise on cytokines and careful review of the manuscript. We thank William Gahl, M.D., Ph.D. for reviewing the manuscript and providing key editing suggestions. Finally, we thank Teresa Mangone, M.D. and Allen Eghrari, M.D. for reviewing the manuscript and providing integral insight from a cornea perspective.

Funding information Supported by the Intramural Research Program of the National Institutes of Health.

Compliance with ethical standards

Conflict of interest The authors declare no competing financial interests.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Hessen M, Akpek EK. Ocular graft-versus-host disease. Curr Opin Allergy Clin Immunol. 2012;12:540–7.
2. Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. Lancet. 2009;373:1550–61.
