



Published in final edited form as:

Bone Marrow Transplant. 2014 January ; 49(1): 116–121. doi:10.1038/bmt.2013.137.

Validation of the National Institutes of Health chronic GVHD Oral Mucosal Score using component-specific measures

CW Bassim^{1,9}, H Fassil^{2,3,9}, JW Mays¹, D Edwards¹, K Baird⁴, SM Steinberg⁵, KM Williams², EW Cowen⁶, SA Mitchell⁷, K Cole², T Taylor², D Avila², D Zhang⁵, D Pulanic^{2,8}, L Grkovic^{2,8}, D Fowler², RE Gress², and SZ Pavletic²

¹National Institute of Dental and Craniofacial Research, NIH, Bethesda, MD, USA

²Experimental Transplantation and Immunology Branch, National Cancer Institute, NIH, Bethesda, MD, USA

³Tufts University School of Dental Medicine, Boston, MA, USA

⁴Pediatric Oncology Branch National Cancer Institute, NIH, Bethesda, MD, USA

⁵Biostatistics and Data Management Section, Office of the Clinical Director, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD, USA

⁶Dermatology Branch National Cancer Institute, NIH, Bethesda, MD, USA

⁷Outcomes Research Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, NIH, Bethesda, MD, USA

⁸Division of Hematology, Department of Internal Medicine, Clinical Hospital Center Zagreb, Zagreb, Croatia

Abstract

Oral chronic GVHD (cGVHD) is a common, late complication of alloSCT that is associated with significant patient morbidity. The NIH Oral Mucosal Score (NIH OMS) was developed to assess oral cGVHD therapeutic response, but has not been fully validated. This study's purpose was to conduct a rigorous construct validity and internal consistency analysis of this score and its components (erythema, lichenoid, ulcers, mucocelles) using established measures of oral pain, oral function, oral-related quality-of-life, nutrition and laboratory parameters in 198 patients with cGVHD. The construct validity of the NIH OMS was supported: a moderate correlation was observed between NIH OMS and mouth pain ($\rho = 0.43$), while a weaker correlation was observed with low albumin ($\rho = -0.26$). Total NIH OMS, erythema and lichenoid components were associated with malnutrition, oral pain and impaired oral QOL, while ulcers were only associated with oral pain. No associations were found between mucocelles and any indicator evaluated, including salivary function or xerostomia. Kappa determined between scale components was low overall (all < 0.35), supporting a conclusion that each component measures a

Correspondence: Dr CW Bassim, National Institute of Dental and Craniofacial Research, Clinical Research Center, NIH, Building 10, Room 1N-117, Bethesda, MD 20892, USA. carol.bassim@nih.gov.

⁹These authors contributed equally to this work.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

distinct manifestation of oral cGVHD. This study supports the use of the NIH OMS and its components (erythema, lichenoid and ulcerations) to measure clinician-reported severity of oral cGVHD.

Keywords

oral cGVHD; chronic graft-versus-host disease; validation

INTRODUCTION

Chronic GVHD (cGVHD) is a major late complication of allogeneic hematopoietic SCT (alloHSCT).¹ It is a clinical syndrome characterized by complex allogeneic and autoimmune dys-regulation of the immune system and is the leading cause of non-relapse-related morbidity and mortality of long-term transplant survivors.² Chronic GVHD may persist for months or years and may affect multiple organ systems including the eyes, mouth, gut, liver, lungs, joints and genitourinary tract. The mouth is commonly involved, with oral manifestations occurring in 45–83% of cGVHD patients.³

Oral manifestations of cGVHD can be characterized as mucosal, salivary and/or sclerotic in nature and resemble several auto-immune conditions including Sjögren's syndrome, oral lichen planus and scleroderma in both clinical features and histological appearance.⁴ The spectrum of clinical presentation of oral cGVHD is diverse and includes erythema, lichenoid hyperkeratosis, xerostomia, mucocles, atrophy, edema, fibrosis, pseudomembrane and ulcerations and can involve any site in the oral cavity. Oral cGVHD can be a significant contributor to pain and discomfort, resulting in diminished oral health, impaired oral cavity function and reduced quality-of-life (QOL).^{5–9} Despite its prevalence and impact on health and well-being, there remains no standard treatment of oral cGVHD.¹⁰

A major obstacle in advancing the development of new therapies for oral GVHD is the lack of well-validated measures to evaluate treatment response in clinical trials. In 2006, the NIH Consensus Development Project published criteria for the measurement of therapeutic response in clinical trials of cGVHD,¹¹ presenting the NIH cGVHD Oral Mucosal Score (NIH OMS) as a measure of oral disease severity. The NIH OMS is a clinician-evaluated measure of the mucosal manifestations of oral cGVHD, developed to increase objectivity and quantification in serial monitoring of oral cGVHD. However, this 15-point proposed scoring system was based on recommendations from a collaborative team of clinicians and required validation in patients with cGVHD. Subsequent studies assessing the validity and reliability of the NIH OMS have often been limited in their study design, size or scope.^{6,12–14} Recently, Treister *et al.*¹⁵ published a large prospective study through the Chronic GVHD Consortium describing the NIH OMS and its components in a cohort of patients with cGVHD and analyzing these findings based on their associations with changes in oral pain. However, the score itself and its components have not yet been fully validated.

In the current study, we assessed the construct validity of the NIH OMS and its components through comparisons with a comprehensive list of accepted measures used in the evaluation of cGVHD and in oral disease using a large cohort of systematically evaluated patients with

cGVHD. The aim was to determine clinical relevance of the NIH OMS and its components and to recommend eventual refinements and simplification of the scoring.

MATERIALS AND METHODS

Study population

Two hundred and sixty-seven post-alloHSCT patients (247 adults, 20 pediatric patients <18 years), referred for evaluation of cGVHD, were enrolled in a prospective cross-sectional study of cGVHD at the NIH Clinical Center in Bethesda, Maryland from 2004 to 2012 (clinicaltrials.gov identifier: NCT00331968). Patients with inconclusive cGVHD ($N=11$) or late acute GVHD ($N=2$), or patients enrolled before the NIH consensus conference or who were not evaluated using the 2006 NIH OMS ($N=56$), were excluded from the study, thus leaving 198 evaluable participants (191 adults, 7 pediatric patients) with cGVHD and with NIH OMS available. This research has been approved by the NCI Center for Cancer Research Institutional Review Board (IRB).

Measures

NIH Oral Mucosal Score—Clinicians at the NIH, advanced practitioners or hematology-oncology fellows experienced in assessing chronic GVHD patients, scored oral cGVHD manifestations using the 15-point oral cGVHD OMS (NIH OMS). This scale evaluates the four most common manifestations of oral cGVHD: erythema, lichenoid lesions, ulcers and mucoceles. It provides a value for the severity of each of these manifestation (erythema, lichenoid, mucoceles scored 0–3, ulcers scored 0–6) as well as a total score (scored 0–15) (Figure 1).¹¹

Patient-reported outcome measures—In this study, we employed data derived from four self-reported measures. The Lee cGVHD Symptom Scale, including a subscale for eyes and mouth symptoms,¹⁶ the NIH Oral Symptom Scores (mouth dryness, oral pain and oral sensitivity on a 0–10 scale),¹¹ the Oral Health Impact Profile (OHIP),¹⁷ (<http://www.asbmt.org/displaycommon.cfm?an=1&subarticlenbr=29>) and the Functional Assessment of Cancer Therapy-Bone Marrow Transplant (FACT-BMT).¹⁸

Nutritional Assessment Scale—Nutritional status was evaluated by a certified dietitian using the patient-generated subjective global assessment (PG-SGA), and was categorized as well-nourished, moderately malnourished or at risk for nutritional deficits or severely malnourished.¹⁹

Laboratory markers of inflammation—Laboratory markers of inflammation total platelets count, C reactive protein, total complement and low albumin were analyzed for associations with NIH OMS.

Salivary flow rate—Unstimulated whole saliva was collected every 30 s for a total of 5 min by dental clinicians. The mass of the collected saliva was determined and then divided by 5 min to determine the 5-min salivary flow rate (28). At the time of our analysis, the salivary flow rate data was available for 87 patients.

NIH organ-specific, global and average scores—NIH organ-specific scores are based on a transplant clinician-reported scale of 0–3 to evaluate each of eight organ systems for women and seven for men: skin, eye, mouth, lung, liver, gastrointestinal tract, joint/fascia and genital tract (women only). The NIH global score (none, mild, moderate, severe) and the NIH average score were also assessed.⁴

Survival—Survival status was determined through phone calls to patients or the offices of primary care providers and searches of the Social Security Death Index.

Statistical analysis

Construct validity—Construct validity, the extent to which a measure is associated in theoretically expected directions with measures of both related and unrelated constructs, was examined by comparing the NIH OMS and its components to a set of conceptually related measures. Relationships between these measures and the total OMS as well as the four constituent components (erythema, lichenoid, ulcers and mucoceles) were explored. The total scale was considered as a continuous parameter (range 0–15) while the components were considered as ordered categorical or dichotomous variables. To accommodate the skewed distribution and the ordinal nature of the variables, non-parametric tests were used. Specifically, comparisons of ordered categorical parameters vs a dichotomous classification variable were evaluated with a Cochran-Armitage trend test.²⁰ Parameters that were both dichotomous were compared using Fisher's exact test. An exact Wilcoxon rank sum test was used to determine the significance of the difference between two groups with respect to a continuous outcome. The Jonckheere–Terpstra test was used to determine the association between two ordered categorical parameters, or between an ordered categorical parameter and a continuous parameter.²¹ An exact Kruskal–Wallis test was used to determine the significance of a continuous parameter evaluated over three or more unordered categories. The association between a dichotomous and an unordered categorical parameter was determined by Mehta's modification to Fisher's exact test.²² Spearman rank correlation was used to determine the correlation between two continuous parameters. For the purposes of this study, $|\rho| > 0.50$ would indicate a moderate to strong correlation, $0.3 < |\rho| < 0.5$ would indicate a weak to moderate correlation and $|\rho| < 0.3$ would indicate weak correlation.²³

Internal consistency—Internal consistency was determined by assessing the relationship between the overall scale and each of its individual components, and the degree to which these components are associated. This was done using the Jonckheere–Terpstra test for trend as the intention was to demonstrate the association between pairs of ordered categorical parameters. When the components were reduced to dichotomous categories, a Kappa statistic was used to evaluate the degree of agreement between two measures.

Survival analyses—The association between OMS and components and overall survival were assessed using Kaplan–Meier curves, beginning at the date the patient enrolled on the NIH natural history study until the date of death or last follow-up. The significance of the difference among a set of Kaplan–Meier curves was determined by a log-rank test.

All *P*-values are two-tailed, and were not formally adjusted to account for multiple comparisons, except for survival analysis as described. However, in view of the number of statistical tests performed, only *P*-values <0.01 were considered to be statistically significant.

RESULTS

Patient demographic, transplant and cGVHD characteristics

A total of 198 patients met the diagnostic criteria for cGVHD as outlined by the NIH cGVHD Consensus Conference on diagnosis and staging.⁴ Table 1 details patient demographics and transplant characteristics. No associations were found between these characteristics and the NIH OMS or its components.

Out of total 198 patients, 145 (73%) patients scored above a NIH OMS =0, and 48 (24%) scored above NIH OMS =2. Of the 145 patients manifesting oral cGVHD, 77% manifested erythema, 75% lichenoid changes, 23% ulcerations and 10% mucoceles. Figure 2 shows the distribution of these component scores across the spectrum of NIH OMS.

Construct validity

Erythema—The score for erythema was significantly associated with Lee Symptom subscales for eyes and mouth (*P* =0.005) and eating and digestion (*P* =0.006) as well as with single-item patient-reported questions from the Lee Symptom Scale, OHIP and other questions concerning mouth pain and ability to eat (Table 2). Nutritional assessment (*P* =0.006), NIH 0–3 organ-specific score for the mouth (*P* <0.0001), number of organs involved (*P* =0.002), NIH average score (*P* =0.006) and total albumin (*P* =0.0004) were also associated with erythema scores.

Salivary flow rate (*P* =0.59) was not associated with erythema, nor was patient-reported oral dryness (*P* =0.02). Neither was survival associated with erythema (*P* =0.16).

Lichenoid—Lichenoid changes were significantly associated with responses to single-item patient-reported questions concerning mouth pain and ability to eat, as well as with the total OHIP score (*P* =0.010) (Table 2). As with erythema, nutritional assessment (*P* =0.005), NIH organ-specific score for the mouth (*P* <0.0001), number of organs involved (*P* =0.004) and total albumin (*P* =0.003) were also associated with lichenoid scores. Salivary flow rate (*P* =0.78), self-report of oral dryness (*P* =0.024) and survival (*P* =0.24) were also not associated with lichenoid findings.

Ulceration—Ulcerative changes were significantly associated with specific single-item Lee symptom and OHIP patient-reported questions about mouth pain (Table 2). The NIH organ-specific score for the mouth (*P* <0.0001) was also associated with ulceration scores. Salivary flow rate (*P* =0.94), oral dryness (*P* =0.21) and survival (*P* =0.38) were not associated with ulcerations.

Mucocele—Mucoceles were not associated with any of the measurements or parameters evaluated.

NIH OMS total score—A weak-to-moderately strong, positive correlation was observed between total NIH OAS and patient-reported mouth pain ($\rho = 0.43$) and mouth sensitivity ($\rho = 0.35$). Further, a weak, negative correlation was seen with albumin levels ($\rho = -0.26$), and a weak, positive correlation was seen with the number of cGVHD organs involved ($\rho = 0.27$). Total scores were also significantly associated with nutritional assessment ($P = 0.001$), and Lee Symptom and OHIP questions concerning mouth pain (Table 2). Salivary flow rate ($P = 0.84$) and survival ($P = 0.23$) were not associated with total score; self-reported mouth dryness was weakly associated with total score ($\rho = 0.23$).

The NIH OMS was also significantly associated with the NIH 0–3 Mouth Score ($P < 0.0001$). There was a strong trend for the NIH OMS to increase with increased NIH Mouth Score, with medians ranging from 0 to 7 as the NIH Mouth Score increases from 0 to 3 (Figure 3).

Internal consistency

Table 3 shows that NIH OMSs were significantly associated with each of the components of the scale: erythema ($P < 0.0001$), lichenoid changes ($P < 0.0001$), ulcerations ($P < 0.0001$) and mucocelles ($P = 0.0003$). Erythema had weak agreement with lichenoid changes ($k = 0.35$), poor agreement with ulcerations ($k = 0.19$) and very poor agreement with mucocelles ($k = 0.03$). Lichenoid changes had poor agreement with ulcerations ($k = 0.16$) and very poor agreement with mucocelles ($k = 0.07$). Mucocelles had very poor agreement with all of the other parameters of the scale.

DISCUSSION

This study reports a comprehensive construct validity evaluation of the NIH OMS as a measure of oral chronic GVHD severity, which is based exclusively on an oral exam performed by the transplant clinician. In prior studies, the NIH OMS had a median inter-rater reliability of 0.7, has been preliminarily validated, and was shown to be feasible and practical for use.^{13,14,24} It has also recently been prospectively validated with the perception of improvement of clinician- and patient-reported measures or oral pain.¹⁵ However, the validation of the NIH OMS and its components against component-specific indicators of disease activity and severity has not been done.

Our findings support the construct validity of the individual components of the NIH OMS in that erythema and lichenoid findings were strongly associated with oral pain and poor oral QOL, function and nutrition, as well as with low serum albumin levels and other indicators of worse systemic cGVHD, while ulcers were associated primarily with oral pain. Though frequently seen in cGVHD patients and a clear manifestation of oral cGVHD, ulceration can also be considered a non-specific finding of oral inflammation or infection.²⁵

The distribution of component score contribution to the total score across the spectrum of NIH OMSs (Figure 2) shows that these contributions are not dominated by any component as oral cGVHD worsens, even though ulcerations are more heavily weighted. It appears that as oral cGVHD manifestations become progressively more severe, ulcers account for about half of oral cGVHD findings, with the remaining half divided between erythema and lichenoid lesions. The poor agreement seen between ulcers and erythema ($k = 0.19$) and

between ulcers and lichenoid ($k=0.16$) supports the notion that ulcerations are capturing a distinct manifestation of oral cGVHD separate from erythema and lichenoid, and more fully describing the complex presentation of oral cGVHD.^{15,24}

A notable point is the lack of association between any parameter studied and mucocoeles, also found by Treister *et al.*,¹⁵ which may support the removal of this component from the score. As 77% of cGVHD patients report some degree of salivary gland involvement,²⁶ these results may indicate a deficit of the NIH OMS to represent salivary dysfunction. Using saliva-specific measures such as salivary flow rate may serve to supplement the NIH OMS in assessing the salivary impact of oral cGVHD. Whole salivary flow offers an objective measure of salivary gland involvement that is feasible and non-invasive. This approach has been used to evaluate therapeutic response in clinical trials to treat dry mouth in oral cGVHD²⁷ and has been validated for use in Sjögren's syndrome,²⁸ which is clinically and histopathologically similar to salivary oral cGVHD.²⁶

Consistent with the observations of other investigators, we found that the total NIH OMS is highly associated with patient-reported oral pain and sensitivity, a major symptom and indicator of oral cGVHD and a driving force for therapeutic symptom control.²⁴ Squaring the Spearman's rho found for the association between NIH OMS and oral pain ($\rho=0.43$) shows that only 19% of the variance in NIH OMS was explained by oral pain. This may indicate that the NIH OMS captures a wider clinical presentation than painful mouth sores, supported by the widely diverse associations found in the present study.

It is also to be noted that the NIH OMS was not associated with other organ system cGVHD, except for the NIH Genital Score for women ($P=0.04$), or with Lee Score subscales that are not relevant to oral symptoms. This supports the construct validity of the NIH OMS in that there is a clear divergence of association between findings that were theoretically expected to be linked with oral cGVHD and findings which were not. Mucosal cGVHD of the mouth can be readily theoretically associated with mucosal cGVHD of the vulva-vaginal region for women.

Several laboratory markers of inflammation, such as total platelet count, C reactive protein, total complement and low albumin have been associated with cGVHD,²⁹ with albumin and total complement specifically correlating with the severity of oral cGVHD.⁶ Our finding that higher total NIH OMS is associated with lower serum albumin replicates this previous observations of an association between hypoalbuminemia and oral cGVHD⁶ as well as overall cGVHD activity,²⁹ which were derived using a subset of this current cohort.

The NIH OMS is based exclusively on a clinician's examination of the oral mucosa of a post-transplant patient. This is in contrast to the NIH Mouth Score, developed to grade oral cGVHD in a clinical setting and which is based on a clinician's integrated judgment of oral disease signs, patient-reported symptoms and limitation of oral intake. The NIH OMS and the erythema, lichenoid and ulcer components were all significantly associated with the NIH 0–3 Mouth Score, showing a significant trend for increasing NIH OMS with increasing NIH Mouth Score. Future studies need to determine the exact relationship and utility of these two scales in assessment and monitoring of oral chronic GVHD.

The limitations of this study arise from its cross-sectional design. The NIH OMS score is intended to evaluate therapeutic response to treatment, and thus prospectively collected serial data points after therapeutic intervention would add knowledge concerning the predictive validity of the NIH OMS and its responsiveness to change. Furthermore, the participants included in this analysis were predominantly those with severe and therapy-refractory cGVHD, thus providing a sample of patients who were sufficiently affected by oral cGVHD to permit a meaningful evaluation of NIH OMS validity; however, these results may not be generalized to individuals with milder cGVHD manifestations.

In conclusion, the proposed NIH OMS offers a clinically relevant description of oral cGVHD, with the erythema, lichenoid and ulceration components independently being associated with cGVHD relevant signs and/or symptoms. The score could be refined by removing the component for mucocelles, as these were not associated with any disease activity. Also, the score could be complemented by evaluation of salivary function, as it does not capture saliva-specific parameters of disease activity. The NIH OMS demonstrates favorable construct validity as an instrument to assess the severity of oral cGVHD manifestations.

Acknowledgments

H Fassil's research year was made possible through the Clinical Research Training Program, a public-private partnership supported jointly by the NIH and Pfizer Inc., and by the Intramural Research Program of the National Cancer Institute, Center for Cancer Research, National Institute of Health.

References

1. Copelan EA. Hematopoietic stem-cell transplantation. *N Engl J Med.* 2006; 354:1813–1826. [PubMed: 16641398]
2. Baird K, Pavletic SZ. Chronic graft versus host disease. *Curr Opin Hematol.* 2006; 13:426–435. [PubMed: 17053454]
3. Schubert MM, Correa ME. Oral graft-versus-host disease. *Dent Clin North Am.* 2008; 52:79–109. [PubMed: 18154866]
4. Filipovich AH, Weisdorf D, Pavletic S, Socie G, Wingard JR, Lee SJ, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. *Biol Blood Marrow Transplant.* 2005; 11:945–956. [PubMed: 16338616]
5. Imanguli MM, Alevizos I, Brown R, Pavletic SZ, Atkinson JC. Oral graft-versus-host disease. *Oral Dis.* 2008; 14:396–412. [PubMed: 18593456]
6. Fassil H, Bassim C, Mays J, Edwards D, Baird K, Steinberg S, et al. Oral chronic graft-vs-host disease characterization using the NIH scale. *J Dent Res.* 2012; 91:45S–51S. [PubMed: 22699667]
7. Fall-Dickson JM, Mitchell SA, Marden S, Ramsay ES, Guadagnini JP, Wu TX, et al. Oral symptom intensity, health-related quality of life, and correlative salivary cytokines in adult survivors of hematopoietic stem cell transplantation with oral chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2010; 16:948–956. [PubMed: 20139026]
8. Meier JK, Wolff D, Pavletic S, Greinix H, Gosau M, Bertz H, et al. Oral chronic graft-versus-host disease report from the International Consensus Conference on clinical practice in cGVHD. *Clin Oral Invest.* 2011; 15:127–139.
9. Bassim C, Ambatipudi K, Mays J, Edwards D, Swatkoski S, Fassil H, et al. Quantitative salivary proteomic differences in oral chronic graft-versus-host disease. *J Clin Immunol.* 2012; 32:1390–1399. [PubMed: 22806177]

10. Treister N, Duncan C, Cutler C, Lehmann L. How we treat oral chronic graft-versus-host disease. *Blood*. 2012; 120:3407–3418. [PubMed: 22898605]
11. Pavletic SZ, Martin P, Lee SJ, Mitchell S, Jacobsohn D, Cowen EW, 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; 12:252–266. [PubMed: 16503494]
12. Elad S, Zadik Y, Zeevi I, Miyazaki A, de Figueiredo MA, Or R. Oral cancer in patients after hematopoietic stem-cell transplantation: long-term follow-up suggests an increased risk for recurrence. *Transplantation*. 2010; 90:1243–1244. [PubMed: 21119507]
13. Mitchell SA, Jacobsohn D, Thormann Powers KE, Carpenter PA, Flowers ME, Cowen EW, 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*. 2011; 17:1619–1629. [PubMed: 21536143]
14. Treister NS, Stevenson K, Kim H, Woo SB, Soiffer R, Cutler C. Oral chronic graft-versus-host disease scoring using the NIH consensus criteria. *Biol Blood Marrow Transplant*. 2010; 16:108–114. [PubMed: 19772943]
15. Treister N, Chai X, Kurland B, Pavletic S, Weisdorf D, Pidala J, et al. Measurement of oral chronic GVHD: results from the chronic GVHD Consortium. *Bone Marrow Transplant*. 2013; 48:1123–1128. [PubMed: 23353804]
16. 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; 8:444–452. [PubMed: 12234170]
17. Slade G, Spencer A. Development and evaluation of the Oral Health Impact Profile. *Community Dent Health*. 1994; 11:3–11. [PubMed: 8193981]
18. McQuellon RP, Russell GB, Cella DF, Craven BL, Brady M, Bonomi A, et al. Quality of life measurement in bone marrow transplantation: development of the Functional Assessment of Cancer Therapy-Bone Marrow Transplant (FACT-BMT) scale. *Bone Marrow Transplant*. 1997; 19:357–368. [PubMed: 9051246]
19. Bauer J, Capra S, Ferguson M. Use of the scored Patient-Generated Subjective Global Assessment (PG-SGA) as a nutrition assessment tool in patients with cancer. *Eur J Clin Nutr*. 2002; 56:779–785. [PubMed: 12122555]
20. Agresti, A. *Categorical Data Analysis*. John Wiley and Sons, Inc; New York, NY: 1990.
21. Hollander, M.; Wolfe, D. *Nonparametric Statistical Methods*, Second Edition. John Wiley and Sons, Inc; New York: 1999.
22. Mehta C, Patel N. A network algorithm for performing Fisher's exact test in $r \times c$ contingency tables. *J Am Stat Assoc*. 1983; 78:427–434.
23. Cohen, J. *Statistical power analysis for the behavioral sciences*. L Erlbaum Associates; Hillsdale, NJ: 1988.
24. Elad S, Zeevi I, Or R, Resnick IB, Dray L, Shapira MY. Validation of the National Institutes of Health (NIH) scale for oral chronic graft-versus-host disease (cGVHD). *Biol Blood Marrow Transplant*. 2010; 16:62–69. [PubMed: 19733252]
25. Mays J, Fassil H, Edward D, Pavletic S, Bassim C. Oral chronic graft-versus-host disease: Current pathogenesis, therapy, and research. *Oral Dis*. 2012; 19:327–346. [PubMed: 23107104]
26. Imanguli MM, Atkinson JC, Mitchell SA, Avila DN, Bishop RJ, Cowen EW, et al. Salivary gland involvement in chronic graft-versus-host disease: prevalence, clinical significance, and recommendations for evaluation. *Biol Blood Marrow Transplant*. 2010; 16:1362–1369. [PubMed: 20353829]
27. Agha-Hosseini F, Mirzaii-Dizgah I, Ghavamzadeh L, Ghavamzadeh A, Tohidast-Acrad Z. Effect of pilocarpine hydrochloride on unstimulated whole saliva flow rate and composition in patients with chronic graft-versus-host disease (cGVHD). *Bone Marrow Transplant*. 2007; 39:431–434. [PubMed: 17310130]

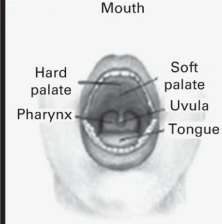
28. Gotoh S, Watanabe Y, Fujibayashi T. Validity of stimulated whole saliva collection as a sialometric evaluation for diagnosing Sjogren's syndrome. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2005; 99:299–302. [PubMed: 15716835]
29. Grkovic L, Baird K, Steinberg SM, Williams KM, Pulanic D, Cowen EW, et al. Clinical laboratory markers of inflammation as determinants of chronic graft-versus-host disease activity and NIH global severity. *Leukemia.* 2012; 26:633–643. [PubMed: 22005783]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



Mucosal change	No evidence of cGVHD		Mild		Moderate		Severe	
	None	0						
Erythema	None	0	Mild erythema or moderate erythema (<25%)	1	Moderate (≥25%) or severe erythema (<25%)	2	Severe erythema (≥25%)	3
Lichenoid	None	0	Hyperkeratotic changes (<25%)	1	Hyperkeratotic changes (25%–50%)	2	Hyperkeratotic changes (>50%)	3
Ulcers	None	0	None	0	Ulcers involving ≤20%	3	Severe ulcers (>20%)	6
Mucoceles*	None	0	1–5 mucoceles	1	5–10 mucoceles	2	Over 10 mucoceles	3

Figure 1.

The NIH OMS (NIH OMS, scored 0–15)¹⁰ measures the severity of each of the most common manifestations of oral cGVHD: erythema, lichenoid, mucoceles (scored 0–3) and ulcers (scored 0–6).

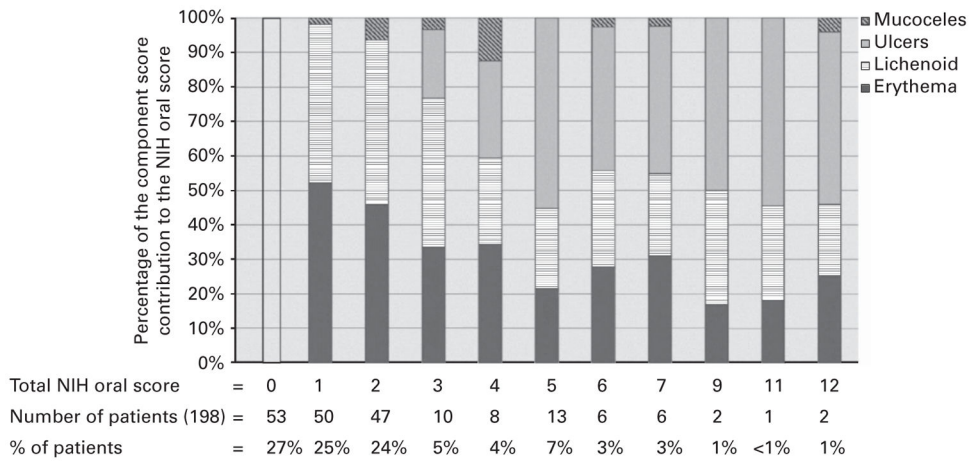


Figure 2. Percentage of component scores to the NIH Oral Score across the spectrum of scores seen, showing the distribution of the components (erythema, lichenoid, ulcers, mucocelles).

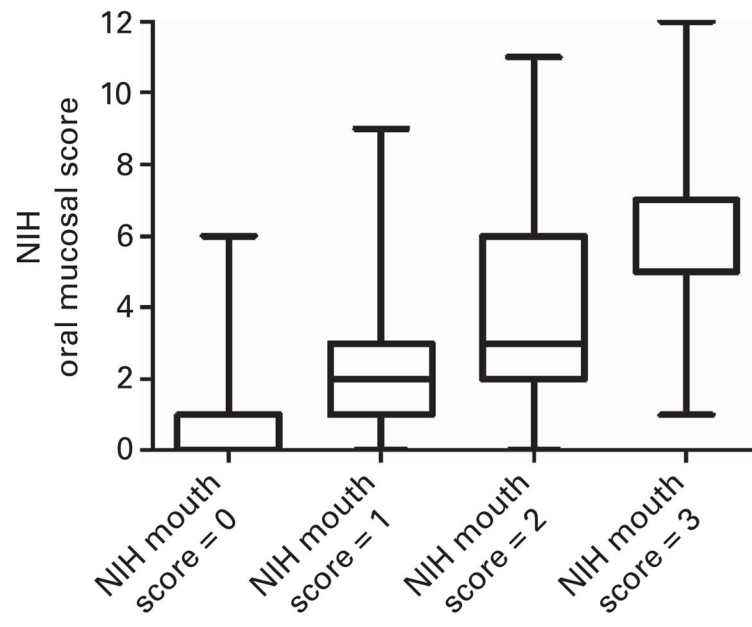


Figure 3. Distribution of NIH mucosal scores across the NIH Mouth Score (boxplot with median and interquartile range shown, whiskers are minimum and maximum values).

Table 1

Patient characteristics at the time of enrollment

Patient characteristics	<i>n</i> (%) or (range)
Total number of patients	198
Age (median, range)	46 (4–64)
<i>Gender</i>	
Male	109 (55%)
Female	89 (45%)
<i>Disease</i>	
ALL/AML/MDS	89 (45%)
CML	29 (15%)
CLL	14 (7%)
HL, NHL	43 (22%)
MM	10 (5%)
Sarcoma	1 (0.5%)
Aplastic anemia/PNH	6 (4%)
Other non-malignant	1 (0.5%)
<i>Conditioning regimen</i>	
Myeloblastic	111 (56%)
Non-myeloblastic	86 (43%)
TBI	78 (39%)
<i>Donor relationship</i>	
Unrelated	75 (36%)
Related	123 (64%)
<i>Gender match: recipient/donor</i>	
Male/male	55 (28%)
Male/female	47 (24%)
Female/female	39 (19%)
Female/male	40 (20%)
Unknown	17 (9%)
<i>Cell source</i>	
Bone marrow	37 (19%)
Peripheral blood	157 (79%)
Cord blood	4 (2%)
<i>HLA match^a</i>	
Yes	160 (81%)
No	32 (16%)
<i>cGVHD onset type</i>	
Progressive	75 (38%)
Quiescent	56 (28%)
<i>de novo</i>	65 (34%)
<i>Activity by therapeutic intent^b</i>	

Patient characteristics	n (%) or (range)
Active	90 (45%)
Not active: decrease systemic therapy	19 (10%)
Not active: cGVHD stable	40 (20%)
Unknown (other)	49 (25%)
<i>Intensity of immunosuppression^c</i>	
None/mild	48 (24%)
Moderate	69 (35%)
Severe	80 (40%)
<i>Number of prior treatments</i>	
<2	22 (11%)
2–5	136 (69%)
>5	34 (17%)
Unknown	4 (2%)
Number of organs involved ^d	5 (1–8)
<i>Individual organs involved^d</i>	
Mouth	135 (68%)
Skin	155 (78%)
Eyes	162 (82%)
Lung	149 (75%)
Liver	102 (51.5%)
Joints or Fascia	124 (63%)
Gastrointestinal tract	93 (47%)
Genitourinary tract	45 (23%)
NIH average score ^e	1.1 (0–2.33)
<i>NIH global score^f</i>	
Mild	3 (2%)
Moderate	59 (30%)
Severe	134 (68%)
Median number of months from transplant to GVHD diagnosis	7 (6–67)
Median number of months from transplant to enrollment	36 (6–223)

Abbreviations: F =female; HL =Hodgkins lymphoma; M =male; MDS = myelodysplastic syndrome; MM =multiple myeloma; NHL =non-Hodgkins lymphoma; PNH =paroxysmal nocturnal hemoglobinuria. For all values in the above table, continuous variables are shown as median values with ranges and categorical variables are shown as frequencies with percentages.

^a HLA match: a minimum of 8/8 allele match in case of unrelated donors and 6/6 antigen and or allele match (HLA-A, -B, -DR) in case of related donors

^b Active: (1) increase systemic therapy because cGVHD is worse; (2) substitute systemic therapy due to lack of response; and (3) withdraw systemic therapy due to lack of response. Not active: (1) decrease systemic therapy because cGVHD is better; (2) not change current systemic therapy because cGVHD is stable; (3) alter systemic therapy owing to its toxicity. Other: either did not receive any immunosuppressive therapy or did not meet any of the criteria.

^c Intensity of immunosuppression: mild, single-agent prednisone<0.5; moderate, prednisone 0.5 mg/kg/day and/or any single agent/modality; high, 2 or more agents/modalities±prednisone 0.5 mg/kg/day.

^d NIH score of 0 not affected versus score>0 affected.

^e NIH average score: total of NIH scores divided by number of organs affected.

^f NIH global score: mild, only 1 or 2 organs (except lung), with max score of 1 in all organs; moderate, at least 1 organ with max score 2 or 3 or more organs with max score of 1 or lung score of 1; severe, at least 1 organ with score of 3 or lung score of 2 or more.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2
Construct validity of NIH Oral Activity Assessment Scale by component and total score^a

Variable	Component (P-value)					Total score	
	Erythema	Lichenoid	Ulcer	Mucocle	Spearman's rho-value ^b	P-value ^c	
<i>Lee: total score</i>	0.015				0.19	—	
Subscale: eyes and mouth	0.005	0.018			0.21	—	
Subscale: eating and digestion	0.006	0.013			NS	—	
Need to avoid certain foods due to mouth pain	< 0.0001	0.0004	< 0.0001		—	< 0.0001	
Ulcers in mouth	0.0002	< 0.0001	< 0.0001		—	< 0.0001	
Difficulty swallowing solid foods	0.004	0.0004			—	0.005	
Difficulty swallowing liquid	0.020				—	—	
<i>OHIP: total score</i>	0.035	0.01			NS	—	
Had painful aching in mouth	0.0005	0.0004	0.011		—	0.0001	
Found it uncomfortable to eat any foods	0.0002	< 0.0001	0.002		—	< 0.0001	
Felt sense of taste has worsened	0.023	0.016			—	0.045	
Have been self-conscious	0.013	0.017			—	0.049	
Had to interrupt meals		0.025			—	—	
Felt life in general was less satisfying	0.046	0.016			—	0.045	
Have been irritable with other people		0.023			—	—	
Oral pain	< 0.0001	< 0.0001	< 0.0001		0.43	—	
Oral sensitivity	0.016	0.0002	0.0002		0.35	—	
Mouth dryness	0.015	0.024			0.23	—	
Albumin	0.0004	0.003			30.26	—	
Total complement	0.025				0.19	—	
Nutritional assessment	0.006	0.005			—	0.0011	
Number of organs involved	0.002	0.004			0.27	—	
NIH average score	0.007	0.050			0.20	—	
NIH Mouth Score	< 0.0001	< 0.0001	< 0.0001		—	< 0.0001	
NIH Genital Score					—	0.04	
FACT-BMT: social/family			0.027		NS	—	

Abbreviations: FACT-BMT =Functional Assessment of Cancer Therapy-Bone Marrow Transplant; NS =non significant.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

^a *P*-values were adjusted for multiple comparisons and are considered significant at $P < 0.01$ (shown in bold); all results found at a $P < 0.05$, however, are provided for this table.

^b Spearman's *r*-values are provided when correlation of two continuous parameters are presented; *P*-values are significant but not provided when the *r*-value is reported.

^c *P*-values are presented for results of measures of association not based on correlation of two continuous parameters.

