



Published in final edited form as:

Transplant Cell Ther. 2021 December ; 27(12): 988.e1–988.e7. doi:10.1016/j.jtct.2021.08.021.

Evaluation of elafin as a prognostic biomarker in acute graft-versus-host disease

Makda Getachew Zewde¹, George Morales¹, Isha Gandhi¹, Umut Özbek², Paibel Aguayo-Hiraldo³, Francis Ayuk⁴, Janna Baez¹, Chantiya Chanswangphuwana⁵, Hannah Choe⁶, Zachariah DeFilipp⁷, Aaron Etra¹, Stephan Grupp⁸, Elizabeth O. Hexner⁹, William Hogan¹⁰, Nora Rebeka Javorniczky¹¹, Stelios Kasikis¹, Carrie L. Kitko¹², Steven Kowalyk¹, Elisabeth Meedt¹³, Pietro Merli¹⁴, Ryotaro Nakamura¹⁵, Muna Qayed¹⁶, Ran Reshef¹⁷, Wolf Rösler¹⁸, Tal Schechter¹⁹, Daniela Weber¹³, Matthias Wölfel²⁰, Gregory Yanik²¹, Rachel Young¹, John E. Levine^{1,*}, James L.M. Ferrara^{1,*}, Yi-Bin Chen^{7,*}

¹Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY

²Biostatistics Shared Resource Facility, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY.

³Children's Center for Cancer and Blood Diseases, Blood and Marrow Transplantation Section, Children's Hospital Los Angeles, Los Angeles, CA

⁴Department of Stem Cell Transplantation, University Medical Center, Hamburg-Eppendorf, Germany

⁵Blood and Marrow Transplantation Program, Chulalongkorn University, Bangkok, Thailand

⁶Blood and Marrow Transplantation Program, Ohio State University, Columbus, OH

⁷Hematopoietic Cell Transplant and Cellular Therapy Program, Massachusetts General Hospital, Boston, MA.

⁸Division of Oncology, The Children's Hospital of Philadelphia, Philadelphia, PA.

⁹Abramson Cancer Center, University of Pennsylvania, Philadelphia, PA

¹⁰Blood and Marrow Transplantation Program, Mayo Clinic, Rochester, MN

¹¹Department of Medicine I, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Germany

¹²Pediatric Blood and Marrow Transplantation Program, Vanderbilt University Medical Center, Nashville, TN

¹³Blood and Marrow Transplantation Program, University of Regensburg, Regensburg, Germany

Corresponding Author: James Ferrara: james.ferrara@mssm.edu, Hematologic Malignancies Translational Research Center, The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, 1 Gustave L. Levy Place/Box 1128, New York, NY 10029.

*These authors contributed equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

¹⁴Department of Pediatric Hematology and Oncology, Bambino Gesù Children's Hospital, Rome, Italy

¹⁵Hematology and Hematopoietic Cell Transplantation, City of Hope Medical Center, Duarte, CA

¹⁶Pediatric Blood and Marrow Transplantation Program, Aflac Cancer and Blood Disorders Center, Emory University and Children's Healthcare of Atlanta, Atlanta, GA

¹⁷Blood and Marrow Transplantation Program, Columbia University, New York, NY

¹⁸Department of Internal Medicine 5, Hematology/Oncology, University Hospital Erlangen, Erlangen, Germany

¹⁹Division of Hematology/Oncology, Department of Pediatrics, University of Toronto, The Hospital for Sick Children, Toronto, ON, Canada

²⁰Pediatric Blood and Marrow Transplantation Program, Children's Hospital, University of Würzburg, Würzburg, Germany

²¹Blood and Marrow Transplantation Program, University of Michigan, Ann Arbor, MI

Abstract

Background: Acute graft-versus-host disease (GVHD) is a major cause of mortality in patients receiving hematopoietic cell transplantation (HCT) for hematologic malignancies. The skin is the most commonly involved organ in GVHD. Elafin, a protease inhibitor overexpressed in inflamed epidermis, was previously identified as a diagnostic biomarker of skin GVHD. However, this finding was restricted to a subset of patients with isolated skin GVHD. The main driver of nonrelapse mortality (NRM) in HCT patients is GI GVHD. Two biomarkers, Regenerating islet-derived 3a (REG3a) and Suppressor of tumorigenesis 2 (ST2), have been validated as biomarkers of GI GVHD that predict long-term outcomes in patients treated for GVHD. We undertook this study to determine the utility of elafin as a prognostic biomarker in the general population of acute GVHD patients in whom GVHD may develop in multiple organs.

Objective: To analyze serum elafin concentrations as a predictive biomarker of acute GVHD outcomes and to compare it to ST2 and REG3a in a large group of patients treated at multiple centers.

Study Design: 526 patients who received corticosteroid treatment for skin GVHD and who had not been previously studied were analyzed from the Mount Sinai Acute GVHD International Consortium (MAGIC). Serum concentrations of elafin, ST2 and REG3a were measured for all patients using ELISA. Patients were divided randomly into equal training and validation sets and a competing risk regression model was developed to model 6-month NRM using elafin concentration in the training set. Additional models were developed using concentrations of ST2 and REG3a, or the combination of all three biomarkers as predictors. ROC curves were constructed using the validation set to evaluate the predictive accuracy of each model and to stratify patients into high- and low-risk biomarker groups. The cumulative incidence of 6-month NRM, overall survival, and four-week treatment response were compared between risk groups.

Results: Patients in the low-risk elafin group unexpectedly demonstrated a higher incidence of 6-month NRM, although this difference was not statistically significant (17% vs. 11%, $P=0.19$).

Overall survival at 6 months (68% vs. 68%, $P>0.99$) and four-week response (78% vs. 78%, $P=0.98$) were similar in the low- and high-risk elafin groups. The area under the receiver operating curve (AUROC) for elafin was 0.55 whereas it was 0.75 for the combination of ST2 and REG3a. The addition of elafin to the other two biomarkers did not improve the AUROC.

Conclusion: Serum elafin concentrations measured at the initiation of systemic treatment for acute GVHD do not predict 6-month NRM, overall survival, or treatment response in a multicenter population of patients treated systemically for acute GVHD. As seen in previous studies, serum concentrations of the GI GVHD biomarkers ST2 and REG3a were significant predictors of NRM and the addition of elafin levels did not improve their accuracy. These results underscore the importance of GI disease in driving NRM in patients who develop acute GVHD.

Introduction

Allogeneic hematopoietic cell transplantation (HCT) is a curative treatment for a number of high-risk hematologic malignancies. A major cause of mortality in patients receiving HCT is acute graft-versus-host disease (GVHD), a multiorgan disorder in which donor immune cells damage healthy host tissues. The skin, gastrointestinal (GI) tract and liver are the target organs for acute GVHD and together they contribute to the overall GVHD severity, which is graded on a scale of I to IV based on clinical symptoms.

The skin is the most commonly involved organ in acute GVHD (1). Skin GVHD is staged 1–4 based on extent and severity of the rash. A prior study from our group showed that elafin, a protease inhibitor overexpressed in inflamed epidermis, correlated with the extent of skin GVHD and mortality. Plasma concentrations of elafin increased significantly in patients with skin GVHD compared to patients without skin GVHD, to those with a non-GVHD rash, and to those with GVHD isolated to the GI tract. In addition, higher elafin levels correlated with increased GVHD skin stage, and nonrelapse mortality (NRM) was also significantly increased in patients with higher elafin concentrations (2). However, the correlation of elafin concentrations with GVHD outcomes was restricted to a subset of patients with isolated skin GVHD that included those treated with topical therapy alone and excluded patients who developed GI GVHD.

GVHD in the GI tract is the main driver of NRM in patients receiving HCT(3). Two serum biomarkers for GI GVHD have been identified: Regenerating islet-derived 3a (REG3a) and Suppressor of tumorigenesis 2 (ST2) (4, 5). These two biomarkers are released into the serum from damaged GI tissue in the intestinal crypts and, thus, the concentrations of these biomarkers reflect the extent of GI crypt damage (6, 7). The Mount Sinai Acute GVHD International Consortium (MAGIC) created a mathematical algorithm that combines the concentrations of these two biomarkers into the MAGIC algorithm probability (MAP), a single value from 0.001 to 0.999 that estimates the likelihood of NRM when measured at the onset of GVHD. Two thresholds divide the MAPs into three distinct risk groups or Ann Arbor scores (Ann Arbor 1–3) (8). The MAP has also been validated as a response biomarker when measured at several timepoints after treatment for acute GVHD (9). Patients with the largest increases in MAP four weeks after treatment initiation had higher rates of

6-month NRM, and changes in MAP more accurately predicted long-term outcomes than changes in clinical manifestations (9).

Although previous findings demonstrated that measuring GI damage was essential to predict GVHD mortality, the utility of combining GI and skin biomarkers has not been conclusively tested. While our prior study demonstrated a significant association between elafin concentration and mortality (2), subsequent studies have produced conflicting results (10–12) and none of these studies assessed the prognostic value of elafin levels in patients with skin GVHD at the initiation of systemic treatment, regardless of involvement of other target organs. We therefore conducted this study to test the hypothesis that the inclusion of elafin levels in an algorithm that contains ST2 and REG3a would more accurately predict outcomes than either elafin or ST2 and REG3a alone in patients with acute GVHD of the skin that received systemic treatment.

Materials and methods

Study design

MAGIC is comprised of 24 international HCT centers that provide clinical information and serum samples from patients at multiple timepoints after HCT. We identified 839 patients in our database who received allogeneic HCT between January 1, 2014 and December 31, 2018 and who were in remission from their underlying disease at the time they started initial treatment with systemic corticosteroids for acute GVHD (Figure 1) and had a serum sample obtained within ± 3 days for cryopreservation. Patients with skin GVHD at the initiation of treatment (N=526) were divided randomly into training (N=263) and validation (N=263) cohorts with equal distributions of key clinical parameters, such as indication for HCT and other GVHD organ involvement (Table 1). The training cohort was used to develop univariable and multivariable models predicting 6-month NRM. The validation cohort was used to compare the predictive ability of these models.

GVHD clinical criteria

Clinical response to treatment and GVHD staging was determined using published guidelines as previously described (8, 13, 14). All centers were trained in the use of MAGIC GVHD staging and grading guidance prior to participation in this study. Skin GVHD was considered present if: 1) an erythematous rash consistent with the typical appearance of GVHD was present; 2) GVHD was favored in the differential diagnosis; and 3) systemic treatment for GVHD was initiated. Biopsies were not required but if performed, histology needed to be consistent with a diagnosis of GVHD. All MAGIC data coordinators received training in GVHD data extraction and passed a detailed examination before entering data into the database. All data were reviewed centrally and aberrant or unusual scenarios were queried. De-identified data were discussed with senior investigators to clarify staging during monthly webinars when appropriate.

Patients were classified as nonresponders if their GVHD symptoms did not improve or progressed after treatment with systemic or topical steroid therapy, if they received additional systemic immunosuppression to treat GVHD, or if they died within the first 4

weeks of treatment. Complete response (CR) was defined as complete resolution of GVHD symptoms in all involved target organs (skin, GI and liver). Partial response (PR) was defined as improvement without resolution in at least one target organ without worsening in any other target organ. All other responses were categorized as non-response (NR).

Biomarker measurement

Serum samples were collected prospectively on IRB approved protocols and biomarkers were measured retrospectively. Elafin ELISA kits were purchased from R&D Systems and measurements were performed according to the manufacturer's protocol. Samples were diluted 1:50 and both samples and standards were run in duplicate. Concentrations were calculated with SoftMax Pro (Molecular Devices). ST2 and REG3a were analyzed by ELISA as previously described(15). The concentrations of elafin and ST2 are reported in picogram per milliliter and the concentrations of REG3a are reported in nanograms per milliliter.

Statistical analysis

Patient characteristics in the training and validation set were compared using the chi-squared test for categorical variables and the Wilcoxon rank-sum test for continuous variables. Elafin concentrations were compared across skin GVHD stages in the full cohort using the Wilcoxon rank sum test with a Bonferroni correction for multiple comparisons. Competing risk regression models were developed to model the hazard of 6-month NRM in the training cohort using biomarker concentrations. Relapse and second HCT were specified as competing risks. The first model used the elafin concentration alone as the predictor. The second model used the two biomarkers that comprise the MAP, ST2 + REG3a, as predictors. The third model combined all three biomarkers, elafin, ST2 and REG3a as predictors. All three models were derived exclusively from patients in this dataset and without any overlap of patients included in prior studies. Receiver operating characteristic (ROC) curves were constructed using the validation set. Area under the ROC curves (AUC) were compared using the DeLong method(16). An optimized threshold was defined for each model by maximizing the product of the sensitivity and the specificity at the threshold; this threshold was used to stratify patients into high risk and low risk groups. The cumulative incidence of 6-month NRM and relapse were measured after initiation of treatment and differences between high and low risk groups were compared using Gray's test(17). Overall survival (OS) was estimated in high and low risk groups using the Kaplan-Meier method and compared by the log-rank test. Differences between response groups were calculated using chi-squared tests. All analyses were performed using R statistical package, version 3.6.2. (R Core Team 2019).

Results

Patient characteristics

We identified 526 patients who received allogeneic HCT between January 1, 2014 and December 31, 2018, who were in remission, and who had a rash when treatment with systemic corticosteroids was started for acute GVHD a median of 28 days after HCT (Figure 1). Patients included in previous published reports were not included in this dataset.

Fourteen percent of patients were pediatric (age < 18). Topical steroid treatment was concomitantly prescribed to 262 (50%) of patients, but only 4 (<1%) received topical treatment prior to initiation of systemic treatment. Patients were divided randomly into training (N=263) and validation (N=263) cohorts with equal distributions of key clinical characteristics between the training set and validation set (Table 1). In this large group of patients elafin concentrations did not significantly differ by the extent of GVHD rash (Figure 2). Age is a risk factor for GVHD (18, 19) and the distribution of body surface area in children, who have larger heads and smaller extremities, is different than in adults. Thus, we performed a subset analysis in the 73 patients less than 18 years of age and confirmed that elafin concentrations did not correlate with the extent of rash in this younger patient subset. (Figure S1).

Algorithm development and performance

We next developed models to predict 6-month NRM based on biomarker concentrations using patients from the training set (Supplemental Table 1). The ROC curves for each model are shown in Figure 3. The first model using elafin concentration alone had a low AUC (0.55). The second model used the two validated biomarkers ST2 and REG3a and produced an AUC of 0.75 that was statistically superior to the first model ($P=0.02$) and very similar to the results of the previously published algorithm using these two biomarkers (8). A third model used all three biomarkers and resulted in an AUC of 0.76 that was not significantly different to the second model ($P=0.10$). Whereas concentrations of ST2 and REG3a each produced higher hazard ratios of 6-month NRM, elafin concentrations did not produce higher hazard ratios of 6-month NRM, either alone or when added to the other biomarkers in these models (Table S1).

Risk stratification for NRM

We next calculated optimal thresholds for each model that stratified patients into low and high risk groups by maximizing both sensitivity and specificity for each model (Table S1). We then used these thresholds to generate cumulative incidences of NRM in validation set patients (Figure 4). The incidence of 6-month NRM of the entire validation set population was 14% (Figure 4A). Elafin concentrations alone produced no difference in NRM, and in fact the NRM was slightly, but not significantly higher in the low concentration group (17% vs. 11%, $P=0.19$, Figure 4B). As expected, ST2 + REG3a divided patients into two groups with highly different 6-month NRMs (6.7% vs. 31%, $P < 0.001$) which did not change with the inclusion of elafin concentrations in the algorithm (Figure 4C,D). Consistent with previous studies, biomarker concentrations did not predict relapse in any of the three models (Figure S3), and thus elafin alone did not predict overall survival (OS) (Figure 4B lower panel). Similar findings were observed in analyses restricted to the subset of pediatric patients (Table S3).

Evaluation of response to treatment at day 28 produced findings similar to the prediction of NRM. 78% of patients experienced complete or partial response (CR/PR) to corticosteroid treatment at 4 weeks (Figure 5A). Elafin concentrations did not predict response at day 28 of treatment, whereas the algorithm that included ST2 and REG3a concentrations did, and the addition of elafin to the algorithm made no significant difference (Figure 5B–D).

Given that previous studies of elafin evaluated its prognostic ability in the context of isolated skin GVHD, we performed a subset analysis to assess each model's performance in patients with skin GVHD who never developed GVHD in other organs. The 6-month NRM in this subset was much lower compared to the full validation set (7.2%), as expected (Figure S3 panel A). In this subset of patients who never experienced GI disease, none of the models successfully stratified patients into risk groups with different NRM (Figure S3 panels B-D).

Discussion

The diagnosis and management of skin GVHD can be challenging. Rashes caused by acute GVHD share visual similarities with rashes caused by drugs, and skin biopsies are difficult to interpret and are often inconclusive (20, 21). Therefore, it is important to note that the criteria used to diagnose and stage skin GVHD in this study were standardized across all centers (13). Even standardized staging criteria do not ensure the accurate prediction of GVHD outcomes using clinical criteria alone, an area in which biomarkers have shown promise but have not always shown consistent results. Prior studies of elafin expression in the epidermis and its concentration in the plasma have shown inconsistent correlation with skin GVHD diagnosis, severity and outcomes (11, 12, 22, 23); but none of these studies assessed the value of adding elafin to GI GVHD biomarkers in patients with skin GVHD regardless of GI involvement at the initiation of systemic treatment. In this large multicenter analysis, we found that serum elafin concentrations in such patients did not predict 6-month NRM, 6-month overall survival, or treatment response at four weeks. In a model derived from patients in this study and who had not been previously included in any studies of the MAP, serum concentrations of GI GVHD biomarkers ST2 and REG3a were significant predictors of NRM and survival. Contrary to our hypothesis, the addition of elafin levels did not improve upon the predictive accuracy of these validated biomarkers. It should be noted that the GI biomarker model generated by this study (model 2) was as accurate as the previously published MAP for the prediction of NRM, even though these models used different weights for the individual biomarkers ST2 and REG3a (Figure S4).

These results underscore the importance of GI disease in driving NRM in patients who develop acute GVHD. Further evidence for this conclusion derives from the finding that 6-month NRM was significantly lower (7.2%) in patients with isolated skin GVHD who never subsequently developed GI GVHD, and that GI biomarkers were not predictive of NRM in this subset of patients, although this is clearly impossible to know when patients present initially.

In contrast to the initial report of elafin as a biomarker of skin GVHD, the present study found no difference in elafin concentration between clinical stages of skin disease and no correlation with 6-month NRM(2). These discrepancies may be explained by several features of the current study: the multicenter nature of the data, the larger number of patients, the paucity of patients with severe stage 4 skin GVHD and, importantly, the exclusion of patients with skin GVHD that did not require systemic treatment. Although elafin concentrations were not significantly higher for stage 4 compared to stage 1–3 skin GVHD, this finding is limited by the small number of patients with stage 4 disease (N=6). We also note that the exclusion of patients whose skin GVHD was managed with topical

therapy alone removed patients with the most favorable outcomes, which may help explain why elafin was not predictive for NRM. We have no data regarding the rationale behind the choice of topical vs systemic treatment for stage 1–2 skin GVHD, but it is possible that patients with more intense erythema, a factor not included in current GVHD staging criteria, received systemic treatment more frequently. If so, the greater skin inflammation in these patients might result in higher elafin levels than those treated with topical therapy. Indeed, the elafin levels in patients with stage 1–2 skin GVHD in this study were three-fold higher than in our prior publication which included patients treated with topical therapy alone (2). Nevertheless we conclude that elafin has poor prognostic value for patients with skin GVHD requiring systemic treatment. Despite these findings, recent studies have suggested that elafin may still have some utility as a diagnostic biomarker (2, 12, 23).

Nearly 80% of all acute GVHD cases involve the skin(24), and patients with isolated skin disease are more likely to respond to treatment and survive compared to those who develop GVHD of the lower GI tract(3, 25). The treatment for acute GVHD, which includes administration of high-dose systemic corticosteroids, is associated with significant morbidity, including higher risk of infections, decreased physical functioning and worse quality of life(26–28). The MAP identifies a large subset of patients with newly diagnosed acute GVHD who have little evidence of GI crypt destruction and who never develop severe disease; such low-risk patients are candidates for novel treatment strategies that reduce or eliminate aggressive steroid treatment and who might be spared its toxicities. Clinical trials are currently in progress to test such hypotheses (NCT03139604), although elafin has no value in identifying such patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by the American Society of Hematology HONORS (Hematology Opportunities for the Next Generation of Research Scientists) award, the Tisch Cancer Institute Summer Scholars Program, the Pediatric Cancer Foundation, and a grant from the National Cancer Institute (P01CA03942).

Financial Disclosure Statement

Drs. Özbek, Ferrara, and Levine are co-inventors on a GVHD biomarker patent and receive royalties from Viracor.

References

1. Zeiser R, Blazar BR. Acute Graft-versus-Host Disease — Biologic Process, Prevention, and Therapy. *New England Journal of Medicine*. 2017;377(22):2167–79.
2. Paczesny S, Braun TM, Levine JE, Hogan J, Crawford J, Coffing B, et al. Elafin Is a Biomarker of Graft-Versus-Host Disease of the Skin. *Science Translational Medicine*. 2010;2(13):13ra2–ra2.
3. Castilla-Llorente C, Martin PJ, McDonald GB, Storer BE, Appelbaum FR, Deeg HJ, et al. Prognostic factors and outcomes of severe gastrointestinal GVHD after allogeneic hematopoietic cell transplantation. *Bone Marrow Transplant*. 2014;49(7):966–71. [PubMed: 24777184]
4. Ferrara JL, Harris AC, Greenson JK, Braun TM, Holler E, Teshima T, et al. Regenerating islet-derived 3-alpha is a biomarker of gastrointestinal graft-versus-host disease. *Blood*. 2011;118(25):6702–8. [PubMed: 21979939]

5. Vander Lugt MT, Braun TM, Hanash S, Ritz J, Ho VT, Antin JH, et al. ST2 as a Marker for Risk of Therapy-Resistant Graft-versus-Host Disease and Death. *New England Journal of Medicine*. 2013;369(6):529–39.
6. Zhao D, Kim YH, Jeong S, Greenson JK, Chaudhry MS, Hoepfing M, et al. Survival signal REG3 α prevents crypt apoptosis to control acute gastrointestinal graft-versus-host disease. *J Clin Invest*. 2018;128(11):4970–9. [PubMed: 30106382]
7. Reichenbach DK, Schwarze V, Matta BM, Tkachev V, Lieberknecht E, Liu Q, et al. The IL-33/ST2 axis augments effector T-cell responses during acute GVHD. *Blood*. 2015;125(20):3183–92. [PubMed: 25814531]
8. Hartwell MJ, Ozbek U, Holler E, Renteria AS, Major-Monfried H, Reddy P, et al. An early-biomarker algorithm predicts lethal graft-versus-host disease and survival. *JCI Insight*. 2017;2(3):e89798. [PubMed: 28194439]
9. Srinagesh HK, Özbek U, Kapoor U, Ayuk F, Aziz M, Ben-David K, et al. The MAGIC algorithm probability is a validated response biomarker of treatment of acute graft-versus-host disease. *Blood Advances*. 2019;3(23):4034–42. [PubMed: 31816061]
10. Levine JE, Braun TM, Harris AC, Holler E, Taylor A, Miller H, et al. A prognostic score for acute graft-versus-host disease based on biomarkers: a multicentre study. *The Lancet Haematology*. 2015;2(1):e21–e9. [PubMed: 26687425]
11. Brüggem M-C, Petzelbauer P, Greinix H, Contassot E, Jankovic D, French L, et al. Epidermal Elafin Expression Is an Indicator of Poor Prognosis in Cutaneous Graft-versus-Host Disease. *Journal of Investigative Dermatology*. 2015;135(4):999–1006.
12. Solán L, Carbonell D, Muñoz P, Dorado N, Landete E, Chicano-Lavilla M, et al. Elafin as a Predictive Biomarker of Acute Skin Graft-Versus-Host Disease After Haploidentical Stem Cell Transplantation Using Post-Transplant High-Dose Cyclophosphamide. *Frontiers in Immunology*. 2021;12(1).
13. Harris AC, Young R, Devine S, Hogan WJ, Ayuk F, Bunworasate U, et al. International, multicenter standardization of acute graft-versus-host disease clinical data collection: A report from the Mount Sinai Acute GVHD International Consortium. *Biol Blood Marrow Transplant*. 2016;22(1):4–10. [PubMed: 26386318]
14. MacMillan ML, DeFor TE, Weisdorf DJ. The best endpoint for acute GVHD treatment trials. *Blood*. 2010;115(26):5412–7. [PubMed: 20388871]
15. Hartwell MJ, Özbek U, Holler E, Renteria AS, Major-Monfried H, Reddy P, et al. An early-biomarker algorithm predicts lethal graft-versus-host disease and survival. *JCI Insight*. 2018;2(3).
16. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–45. [PubMed: 3203132]
17. Gray RJ. A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. *The Annals of Statistics*. 1988;16(3):1141–54.
18. Qayed M, Wang T, Hemmer MT, Spellman S, Arora M, Couriel D, et al. Influence of Age on Acute and Chronic GVHD in Children Undergoing HLA-Identical Sibling Bone Marrow Transplantation for Acute Leukemia: Implications for Prophylaxis. *Biol Blood Marrow Transplant*. 2018;24(3):521–8. [PubMed: 29155316]
19. Lazaryan A, Weisdorf DJ, DeFor T, Brunstein CG, MacMillan ML, Bejanyan N, et al. Risk Factors for Acute and Chronic Graft-versus-Host Disease after Allogeneic Hematopoietic Cell Transplantation with Umbilical Cord Blood and Matched Sibling Donors. *Biol Blood Marrow Transplant*. 2016;22(1):134–40. [PubMed: 26365153]
20. Marra DE, McKee PH, Nghiem P. Tissue eosinophils and the perils of using skin biopsy specimens to distinguish between drug hypersensitivity and cutaneous graft-versus-host disease. *J Am Acad Dermatol*. 2004;51(4):543–6. [PubMed: 15389188]
21. Reshef R, Saber W, Bolaños-Meade J, Chen G, Chen Y-B, Ho VT, et al. Acute GVHD Diagnosis and Adjudication in a Multicenter Trial: A Report From the BMT CTN 1202 Biorepository Study. *Journal of Clinical Oncology*. 2021;39(17):1878–87. [PubMed: 33507810]

22. George L, Mahabal G, Mohanan E, Balasubramanian P, Peter D, Pulimood S, et al. Limited utility of Plasma Elafin as a biomarker for Skin Graft Versus Host Disease following allogeneic stem cell transplantation. *Clin Exp Dermatol*. 2021.
23. Mahabal GD, George L, Peter D, Bindra M, Thomas M, Srivastava A, et al. Utility of tissue elafin as an immunohistochemical marker for diagnosis of acute skin graft-versus-host disease: a pilot study. *Clinical and Experimental Dermatology*. 2019;44(2):161–8. [PubMed: 29882232]
24. Martin PJ, Schoch G, Fisher L, Byers V, Anasetti C, Appelbaum FR, et al. A Retrospective Analysis of Therapy for Acute Graft-Versus-Host Disease: Initial Treatment. *Blood*. 1990;76(8):1464–72. [PubMed: 2207321]
25. MacMillan ML, Robin M, Harris AC, DeFor TE, Martin PJ, Alousi A, et al. A refined risk score for acute graft-versus-host disease that predicts response to initial therapy, survival, and transplant-related mortality. *Biol Blood Marrow Transplant*. 2015;21(4):761–7. [PubMed: 25585275]
26. Deeg HJ. How I treat refractory acute GVHD. *Blood*. 2007;109(10):4119–26. [PubMed: 17234737]
27. Lee SJ, Kim HT, Ho VT, Cutler C, Alyea EP, Soiffer RJ, et al. Quality of life associated with acute and chronic graft-versus-host disease. *Bone Marrow Transplant*. 2006;38(4):305–10. [PubMed: 16819438]
28. Miller HK, Braun TM, Stillwell T, Harris AC, Choi S, Connelly J, et al. Infectious risk after allogeneic hematopoietic cell transplantation complicated by acute graft-versus-host disease. *Biol Blood Marrow Transplant*. 2017;23(3):522–8. [PubMed: 28017733]

Highlights

- Serum elafin concentrations do not predict 6-month NRM, overall survival or treatment response
- Regenerating islet-derived 3a (REG3a) and Suppressor of tumorigenesis 2 (ST2) are validated biomarkers of GI GVHD that accurately predict 6-month NRM, overall survival and treatment response
- Addition of elafin to REG3a and ST2 did not improve their predictive accuracy

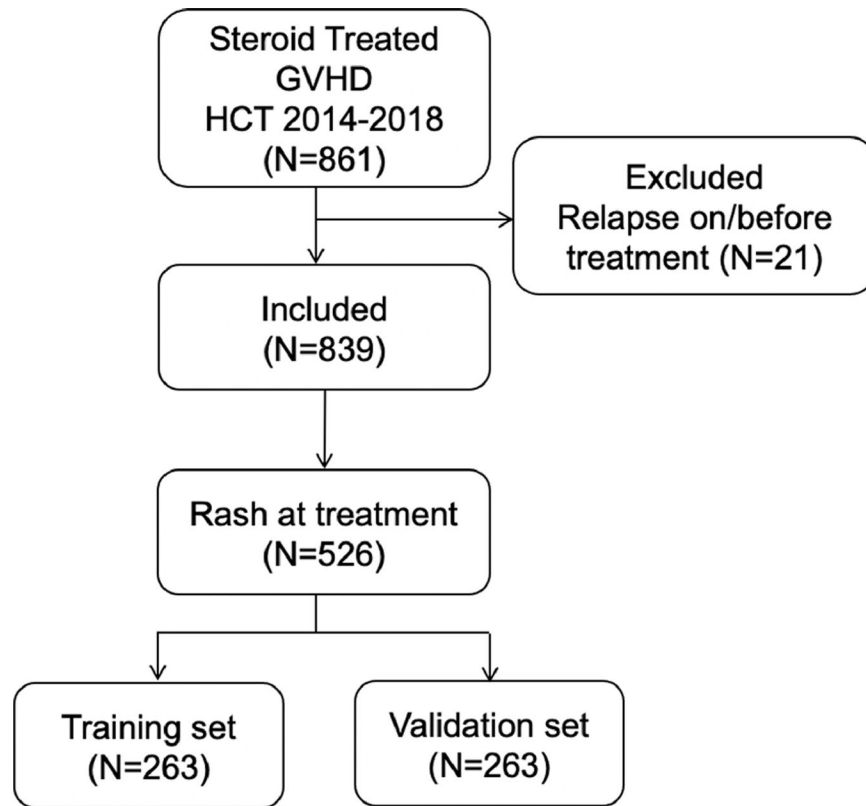


Figure 1. Study flow chart.

Patients who received allogeneic HCT between 1 January 2014 and 31 December 2018, received treatment with systemic corticosteroids following diagnosis of acute GVHD, had serum samples at the time of treatment, and had not been part of the previous validation set were included (N=861). Patients were excluded if they experienced relapse on or before GVHD treatment (N=21). Among these patients, only those who presented with a skin rash at treatment were included (N=526). Patients were randomly assigned to equal training (N=263) and validation (N=263) sets.

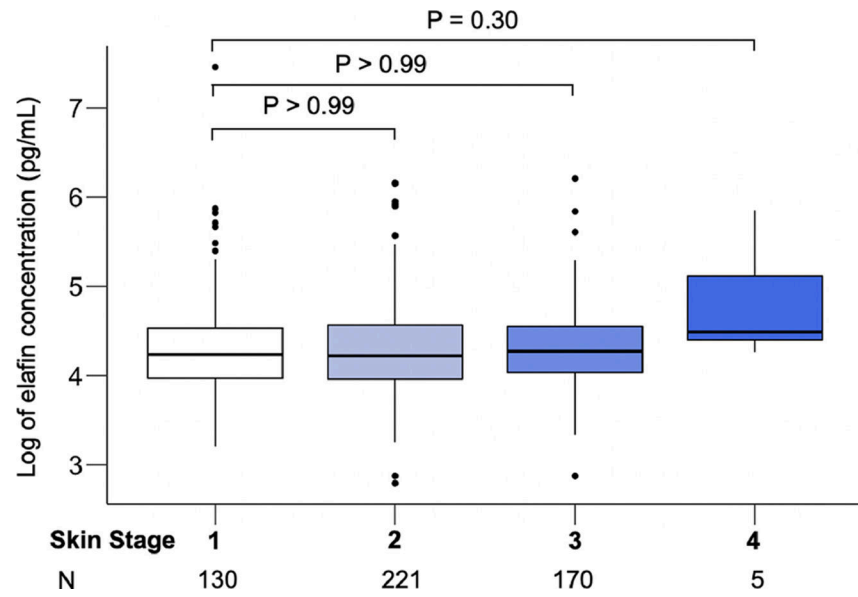


Figure 2. Elafin concentration by skin stage.

Elafin was measured by ELISA and the log of the concentration was correlated with GVHD skin stage (Total N=526). P values were corrected for multiple comparisons.

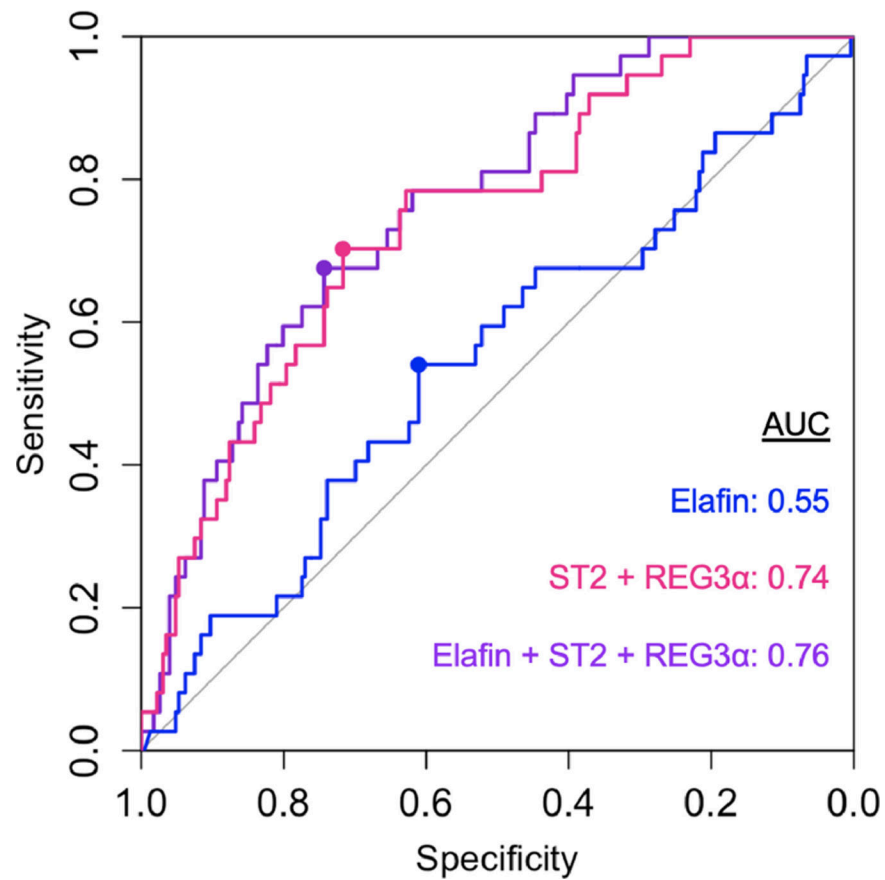


Figure 3. Receiver operating characteristic curve for biomarkers.

Receiver operating characteristic curve for elafin (blue) ST2 + REG3 α (pink) and elafin + ST2 + REG3 α (purple) for prediction of 6-month nonrelapse mortality in patients of the validation set (N=263). The circle indicates the threshold that maximizes the sensitivity and specificity. Threshold between low risk and high risk for elafin = 3.8×10^{-3} , ST2 + REG3 α = 0.16, elafin + ST2 + REG3 α = 3.3×10^{-3} . AUC = area under curve.

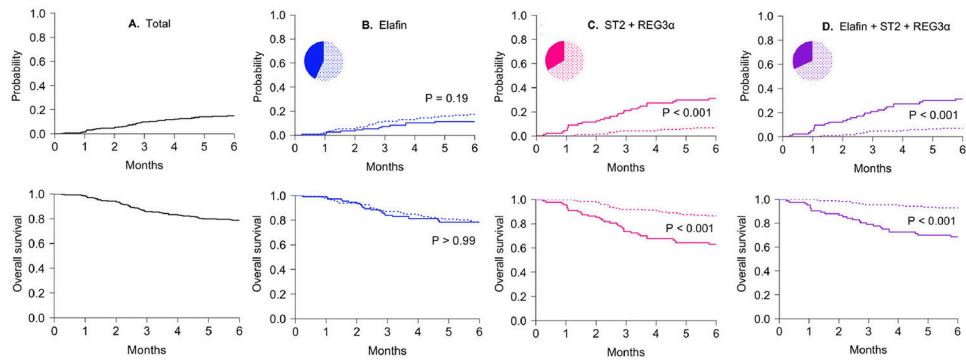


Figure 4. Cumulative incidence of nonrelapse mortality and overall survival in high and low biomarker groups.

Six-month cumulative incidences of nonrelapse mortality (NRM) in high (solid line) and low (dotted line) risk biomarker groups defined by optimized biomarker thresholds and compared using Gray's test (upper panels). Six-month overall survival was estimated using the Kaplan-Meier method and compared by the log-rank test (lower panels). **(A)** Cumulative incidence of NRM (14%) and overall survival (75%) in the total validation set (N=263). **(B)** Cumulative incidence of NRM in the low (N=150) and high (N=113) elafin group (17% vs. 11%, $P=0.19$). Overall survival in the low and high elafin group (68% vs. 68%, $P > 0.99$). Threshold = 3.8×10^{-3} **(C)** Cumulative incidence of NRM in the low (N=175) and high (N=88) ST2 + REG3 α group (6.7 vs. 31%, $P < 0.001$). Overall survival in the low and high ST2 + REG3 α group (77% vs. 51%, $P < 0.001$). Threshold = 0.16 **(D)** Cumulative incidence of NRM in the low (N=180) and high (N=83) elafin + ST2 + REG3 α group (7.0 vs. 30%, $P < 0.001$). Overall survival in the low and high elafin + ST2 + REG3 α group (79% vs. 64%, $P < 0.001$). Threshold = 3.3×10^{-3}

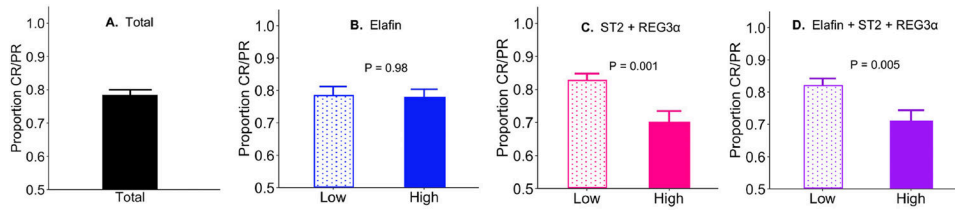


Figure 5. Treatment response for GVHD by high- and low-risk biomarker groups. Proportions of patients who experienced complete or partial response (CR/PR) to corticosteroid treatment at four weeks were compared between high- and low-risk biomarker groups. **(A)** Rate of CR/PR (78%) in the total validation set (N=263). **(B)** CR/PR in the low and high elafin group (78% vs. 78%, $P=0.98$). **(C)** CR/PR in the low and high ST2 + REG3α group (82% vs. 70%, $P=0.002$). **(D)** CR/PR in the low and high elafin + ST2 + REG3α group (82% vs. 71%, $P=0.005$). Error bars represent one standard error of the proportion.

Table 1.

Patient Characteristics (N = 526).

Characteristic	Training set (n=263)	Validation set (n=263)	P
Median age at BMT: yr (range)	54 (0–79)	55 (0–76)	0.97
Age < 18: no. (%)	37 (14)	36 (14)	>0.99
Indication for HCT: no. (%)			0.49
Acute Leukemia	128 (49)	134 (51)	
Lymphoma	27 (10)	25 (10)	
MDS/MPN	75 (29)	69 (26)	
Other Malignant	20 (8)	14 (5)	
Other non-Malignant	13 (5)	21 (8)	
Donor type: no. (%)			0.85
Haploidentical	20 (8)	21 (8)	
Related	50 (19)	45 (17)	
Unrelated	193 (73)	197 (75)	
HLA match: no (%)			0.39
Haploidentical	20 (8)	21 (8)	
Matched	192 (73)	203 (77)	
Mismatched	51 (19)	39 (15)	
Stem cell source: no (%)			0.40
Bone marrow	55 (21)	51 (19)	
Peripheral blood	191 (73)	202 (77)	
Cord blood	16 (6)	10 (4)	
GVHD prophylaxis: no. (%)			0.51
CNI based	206 (83)	210 (84)	
Cyclophosphamide based	40 (16)	35 (14)	
T Cell Depletion	3 (1)	6 (2)	
GVHD serotherapy prophylaxis: no. (%)			
ATG	118 (45)	119 (45)	
No ATG	145 (55)	144 (55)	
GVHD organ involvement at treatment			0.93
Skin only, skin stage 1	38 (14)	37 (14)	
Skin only, skin stage 2	84 (32)	81 (31)	
Skin only, skin stage 3	66 (25)	68 (26)	
Skin only, skin stage 4	2 (1)	2 (1)	
Skin + GI + liver	71 (27)	70 (27)	
Skin + liver	2 (0.8)	5 (1.9)	
Conditioning regimen intensity: no. (%)			>0.99
Full	154 (59)	154 (59)	
Reduced	109 (41)	109 (41)	