



# Optimal Biomechanical Parameters for Measuring Sclerotic Chronic Graft-Versus-Host Disease

Laura X. Baker<sup>1,2</sup>, Fuyao Chen<sup>1,2,3,4</sup>, Austin Cronin<sup>1,2,3</sup>, Heidi Chen<sup>5,6</sup>, Arved Vain<sup>7</sup>, Madan Jagasia<sup>5</sup> and Eric R. Tkaczyk<sup>1,2,3,5</sup>

Skin biomechanical parameters (dynamic stiffness, frequency, relaxation time, creep, and decrement) measured using a myotonometer (MyotonPRO) could inform the management of sclerotic disease. To determine which biomechanical parameter(s) can accurately differentiate patients with sclerotic chronic graft-versus-host disease from post-hematopoietic cell transplant controls, 15 patients with sclerotic chronic graft-versus-host disease and 11 post-hematopoietic cell transplant controls were measured with the myotonometer on 18 anatomic sites. Logistic regression and two machine learning algorithms (least absolute shrinkage and selection operator regression and random forest) were developed to classify subjects. In univariable analysis, frequency had the highest overfit-corrected area under the curve (0.91). Backward stepwise selection and random forest machine learning identified frequency and relaxation time as the optimal parameters for differentiating patients with sclerotic chronic graft-versus-host disease from post-hematopoietic cell transplant controls. Least absolute shrinkage and selection operator regression selected the combination of frequency and relaxation time (overfit-corrected area under the curve = 0.87). Discriminatory ability was maintained when only the sites accessible while the patient is supine (12 sites) were used. We report the distribution of values for these highly discriminative biomechanical parameters, which could inform the assessment of disease severity in future quantitative biomechanical studies of sclerotic chronic graft-versus-host disease.

*JID Innovations* (2021);1:100037 doi:10.1016/j.xjidi.2021.100037

## INTRODUCTION

Chronic graft-versus-host disease (cGVHD) occurs in 30–40% of patients after allogeneic hematopoietic cell transplantation (HCT) and is the leading cause of nonrelapse mortality in patients surviving >2 years after transplantation (Arai et al., 2015; Socié et al., 1999). Among patients with cGVHD, 20% develop sclerosis within 3 years of initial systemic treatment for cGVHD, resulting in significant disability (Inamoto et al., 2013). However, the lack of quantitative and objective

measurements of skin activity has restricted the advancement of treatment options for sclerotic disease (Socié and Ritz, 2014).

The MyotonPRO, a handheld myotonometer that measures soft tissue biomechanical parameters through a noninvasive mechanical impulse, has promise in sclerotic cGVHD management (Baker et al., 2021; Dellalana et al., 2019; Vain, 2012). The device simultaneously extracts five biomechanical parameters: dynamic stiffness, oscillation frequency, relaxation time, creep, and decrement. It has shown high interoperator reliability to measure dynamic stiffness in healthy controls and in patients with sclerotic cGVHD (Chen et al., 2019b; Dellalana et al., 2019). Although a previous study showed the dynamic stiffness parameter's ability to differentiate patients with sclerosis from healthy volunteers, it did not examine other available parameters and lacked post-HCT controls (Chen et al., 2019a).

With the advent of new diagnostic technologies and quantitative methods, an effective coupling of data to statistical methods, including machine learning, could enhance informed decision making. Application of traditional regression models and machine learning algorithms in the setting of skin biomechanical parameters has the potential to differentiate patients with sclerotic cGVHD from post-HCT controls with greater accuracy.

In this cross-sectional study of myotonometry measurements in patients with cGVHD and in post-HCT controls, we investigated the diagnostic ability of all the five available biomechanical parameters in univariable analysis and explored the diagnostic impact of combinations of parameters selected by stepwise regression and machine learning

<sup>1</sup>Dermatology Service and Research Service, Tennessee Valley Healthcare System, U.S. Department of Veterans Affairs, Nashville, Tennessee, USA; <sup>2</sup>Vanderbilt Dermatology Translational Research Clinic, Vanderbilt University Medical Center, Nashville, Tennessee, USA; <sup>3</sup>Department of Biomedical Engineering, School of Engineering, Vanderbilt University, Nashville, Tennessee, USA; <sup>4</sup>Yale School of Medicine, Yale University, New Haven, Connecticut, USA; <sup>5</sup>Vanderbilt-Ingram Cancer Center, Nashville, Tennessee, USA; <sup>6</sup>Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee, USA; and <sup>7</sup>Institute of Physics, University of Tartu, Tartu, Estonia

Correspondence: Eric R. Tkaczyk, Dermatology Service and Research Service, Tennessee Valley Healthcare System, U.S. Department of Veterans Affairs, 719 Thompson Lane, One Hundred Oaks Suite 26300, Nashville, Tennessee 37204, USA. E-mail: eric.tkaczyk@vumc.org

Abbreviations: AIC, Akaike information criterion; AUC, area under the curve; cGVHD, chronic graft-versus-host disease; HCT, hematopoietic cell transplantation; LASSO, least absolute shrinkage and selection operator

Received 16 October 2020; revised 4 May 2021; accepted 12 May 2021; accepted manuscript published online 24 June 2021; corrected proof published online 2 September 2021

Cite this article as: *JID Innovations* 2021;1:100037

algorithms. We also evaluated whether a convenient supine-only protocol maintains the diagnostic accuracy of total body measurements.

## RESULTS

A total of 15 patients with sclerotic cGVHD and 11 post-HCT controls were included in this study. Patient characteristics are summarized in Tables 1 and 2. The median time from HCT was 1,686 days (1–3 quartile: 1,169–2,575 days) for patients with sclerotic cGVHD and 910 days (1–3 quartile: 335–2,507 days) for controls. Subjects were measured with a myotonometer, the MyotonPRO, on 18 anatomic sites (Figure 1). Two patients with cGVHD were measured only on the 12 sites that could be measured with the patient in the supine position owing to discomfort when laying in the prone position. The distribution of parameter values for each anatomic site is shown in Figure 2. Correlation across bilateral sites was high for all the five biomechanical parameters (Table 3). For each subject, we calculated the parameter averages over all the 18 measured sites and over the 12 supine sites. For the averages across all the 18 measured sites, patients with sclerotic cGVHD displayed significant increases in frequency ( $P = 0.033$ ) and stiffness ( $P = 0.043$ ) and decreases in relaxation time ( $P = 0.001$ ) and creep ( $P = 0.002$ ) compared with post-HCT controls (Figure 3). For the averages across the 12 supine sites, patients with sclerotic cGVHD had significant increases in frequency ( $P = 0.033$ ) and decreases in relaxation time ( $P = 0.001$ ) and creep ( $P = 0.002$ ).

### Univariable analysis

For each parameter, averages across all sites and averages across supine sites were used to develop univariable logistic regressions that modeled the diagnosis of sclerotic cGVHD as a dichotomous variable. For parameter averages over all the 18 measured sites, the univariable overfit-corrected areas under curve (AUCs) for frequency, relaxation time, stiffness, and creep were 0.91, 0.89, 0.88, and 0.86, respectively (Table 4). For parameter averages across the 12 supine sites, the overfit-corrected AUCs for frequency, relaxation time, stiffness, and creep were 0.87, 0.87, 0.84, and 0.76, respectively.

### Backward stepwise selection

Backward stepwise selection using the Akaike information criterion (AIC) was applied to a starting multivariable logistic regression model that consisted of all the five biomechanical parameters. The final model for averages across all the sites consisted of the single parameter frequency, whereas the final model for averages across the supine sites consisted of the single parameter relaxation time.

### Machine learning models and bivariable logistic regression

Regression with least absolute shrinkage and selection operator (LASSO) is a machine learning algorithm that performs autonomous variable selection to create reduced models, which can include either a single parameter or a combination of parameters. For both averages across all the sites and averages across the supine sites, LASSO regression selected the combination of frequency and relaxation time as the most important variables in differentiating patients with sclerotic cGVHD from post-HCT controls (overfit-corrected

AUC = 0.87 for averages across all the sites and 0.78 for averages across the supine sites) (Table 5).

A random forest classification algorithm ranked the five biomechanical parameters by their importance in differentiating patients with sclerotic cGVHD from post-HCT controls. Ten of 10 iterations (Table 6) identified frequency as the most important variable, followed by relaxation time, consistent with both logistic regression and LASSO regression models.

The overfit-corrected AUCs from bivariable logistic regressions of frequency and relaxation time for averages across all the sites and averages across the supine sites were 0.89 and 0.85, respectively (Table 4).

### Analysis of anatomic sites

LASSO regression was performed on each anatomic site. The combinations of parameters determined by LASSO as the most important variables for each site are shown in Table 7. Additional univariable analyses examined the discriminatory ability of each of the five biomechanical parameters for individual anatomic sites (Table 8). Both frequency and relaxation time trended toward a higher discriminatory ability in the shin, upper back, chest, upper arm, and abdomen (Figure 4). Stiffness demonstrated a higher discriminatory ability in the chest, shin, and upper arm, whereas creep had a higher discriminatory ability in the chest and upper back. Decrement had a high discriminatory ability in the dorsal forearm (Table 8).

### Correlation between biomechanical parameters

Spearman's correlation ( $r$ ) between frequency and relaxation time was  $-0.94$  for both averages across all the sites and averages across the supine sites (Table 9). High correlation (absolute value of  $r$  close to 1) between frequency and relaxation time suggests the presence of redundant information provided by the two parameters. Similarly, stiffness was highly correlated with frequency ( $r = 0.95$  for both averages across all the sites and averages across the supine sites) and negatively correlated with relaxation time ( $r = -0.93$  and  $-0.94$  for averages across all the sites and averages across the supine sites, respectively).

## DISCUSSION

Objective measurement of skin sclerosis is an unmet need for many sclerosing diseases, including cGVHD and systemic sclerosis (Carpenter et al., 2015; Odell et al., 2020). The myotonometer is a promising potential response measure instrument, but many questions must be answered before clinical implementation (Table 10). The device has demonstrated high interoperator reliability in measuring stiffness in healthy controls and in patients with sclerotic cGVHD (Chen et al., 2019b; Dellalana et al., 2019). A recent longitudinal study in patients with sclerotic cGVHD showed that stiffness measurements over time correlated with clinically important changes in disease (Baker et al., 2021). Our previously published cross-sectional analysis of myotonometry measurements lacked post-HCT controls and only evaluated the discriminatory ability of the stiffness parameter (Chen et al., 2019a). By investigating all the five available parameters using logistic regression and machine learning models, this study found that measuring frequency

**Table 1. Demographic and Clinical Characteristics of the 15 Patients with Sclerotic cGVHD**

cGVHD Subjects (n = 15)	Subject ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Summary <sup>1</sup>	
Demographics <sup>2</sup>	Age, y	55	52	40	28	72	59	47	47	71	60	14	65	60	35	70	55 (40–65)	
	Race	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	100% C
	Sex	M	M	M	F	M	M	M	F	F	M	M	M	M	M	M	M	80% M
	BMI	28	24	27	14	31	17	23	34	21	20	22	33	18	23	32	23 (20–31)	
Disease characteristics <sup>3</sup>	Disease histology	AML	MDS	MM	HL	ALL	AML	NHL	ALL	AML	MDS	AML	AML	NHL	ALL	CLL	33% AML 20% ALL 13% NHL 13% MDS 7% MM 7% HL 7% CLL	
Transplant characteristics <sup>4</sup>	Transplantation source	BM	PB	PB	PB	PB	PB	PB	BM	PB	PB	PB	PB	PB	PB	—	12/14 PB 2/14 BM	
cGVHD characteristics at study entry <sup>5</sup>	NIH score BSA involvement	3	3	2	3	2	3	1	2	1	2	3	1	2	1	—	4/14 NIH 1 5/14 NIH 2 5/14 NIH 3	
	NIH score skin feature	3	3	3	3	3	3	3	3	2	3	3	2	3	3	2	80% NIH 3 20% NIH 2	
MyotonPRO session	Days from HCT <sup>6</sup>	1,706	4,604	1,084	1,908	1,169	2,030	2,575	4,429	1,256	1,119	1,339	3,660	1,061	1,686	2,132	1,706 (1,169–2,575)	

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BM, bone marrow; BMI, body mass index; BSA, body surface area; C, Caucasian; cGVHD, chronic graft-versus-host disease; CLL, chronic lymphoblastic leukemia; F, female; HCT, hematopoietic cell transplantation; HL, Hodgkin’s lymphoma; ID, identification; M, male; MDS, myelodysplastic disorder; MM, multiple myeloma; NHL, non-Hodgkin’s lymphoma; NIH, National Institutes of Health; PB, peripheral blood.

<sup>1</sup>Values are shown as median (interquartile range) for continuous variables and n (%) for categorical variables.

<sup>2</sup>Demographics: C, M, and F.

<sup>3</sup>Disease characteristics: AML, ALL, CLL, MDS, MM, NHL, and HL.

<sup>4</sup>Transplant characteristics: PB and BM.

<sup>5</sup>NIH Score BSA involvement: 1: 1–18%, 2: 19–50%, and 3: >50%; NIH score skin feature: 2 for superficial sclerosis and 3 for deep sclerosis.

<sup>6</sup>Days from HCT: days between HCT and MyotonPRO measurement session.

**Table 2. Demographic and Clinical Characteristics of the 11 Post-HCT Controls**

Post-HCT Controls (n = 11)	Subject ID	16	17	18	19	20	21	22	23	24	25	26	Summary <sup>1</sup>	
Demographics <sup>2</sup>	Age, y	66	25	71	52	65	59	59	65	62	63	57	62 (57–65)	
	Race	C	C	C	C	C	C	C	C	AAn	C	AAn	82% C 18% AAn	
	Sex	M	F	M	F	M	M	M	M	M	F	M	73% M	
	BMI	26	40	26	43	31	24	27	26	36	45	35	31 (26–40)	
Disease characteristics <sup>3</sup>	Disease histology	MF	AA	AML	MDS	NHL	AML	CLL	MDS	AML	AML	ALL	36% AML 18% MDS 9% ALL 9% MF 9% AA 9% NHL 9% CLL	
	Transplant source	PB	BM	PB	CB	PB	PB	PB	PB	-	PB	PB	8/10 PB 1/10 BM 1/10 CB	
	MyotonPRO session	Days from HCT <sup>5</sup>	329	1,785	1,301	3,276	910	411	2,507	4,046	857	335	148	910 (335–2,507)

Abbreviation: AA, aplastic anemia; AAn, African American; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BM, bone marrow; BMI, body mass index; C, Caucasian; CB, cord blood; CLL, chronic lymphoblastic leukemia; F, female; HCT, hematopoietic cell transplantation; ID, identification; M, male; MDS, myelodysplastic disorder; MF, primary myelofibrosis; NHL, non-Hodgkin’s lymphoma; PB, peripheral blood.

<sup>1</sup>Values are shown as median (interquartile range) for continuous variables and n (%) for categorical variables.

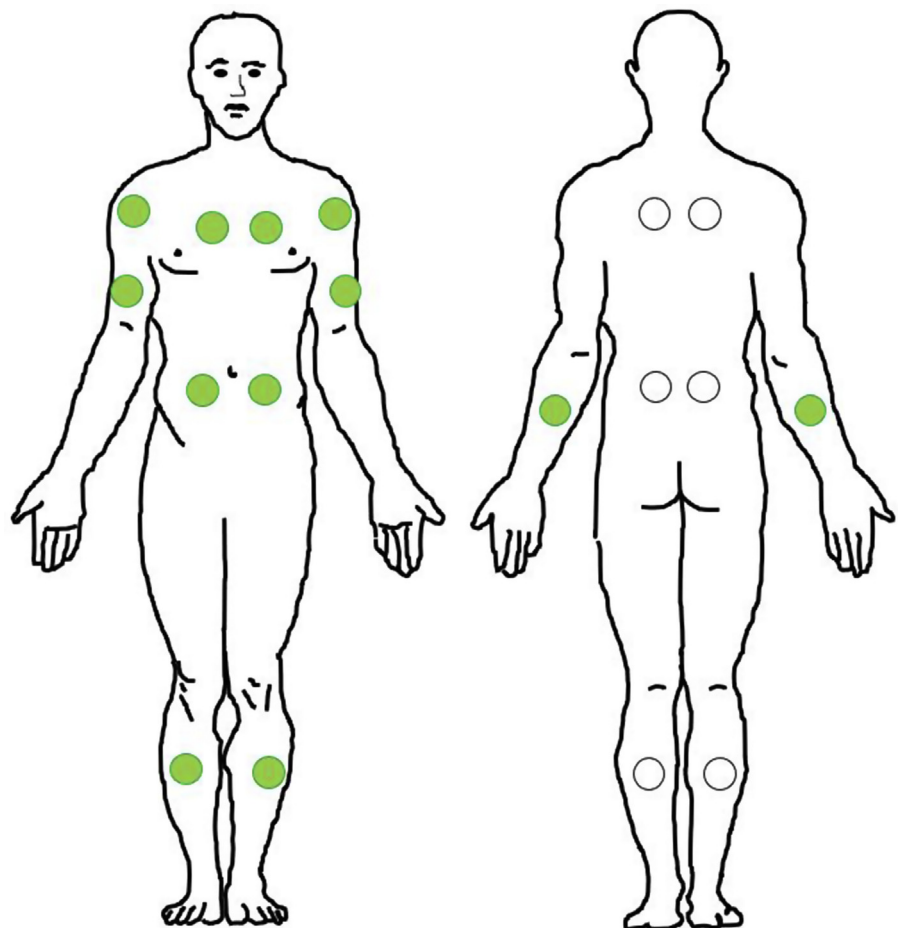
<sup>2</sup>Demographics: C, AAn, M, and F.

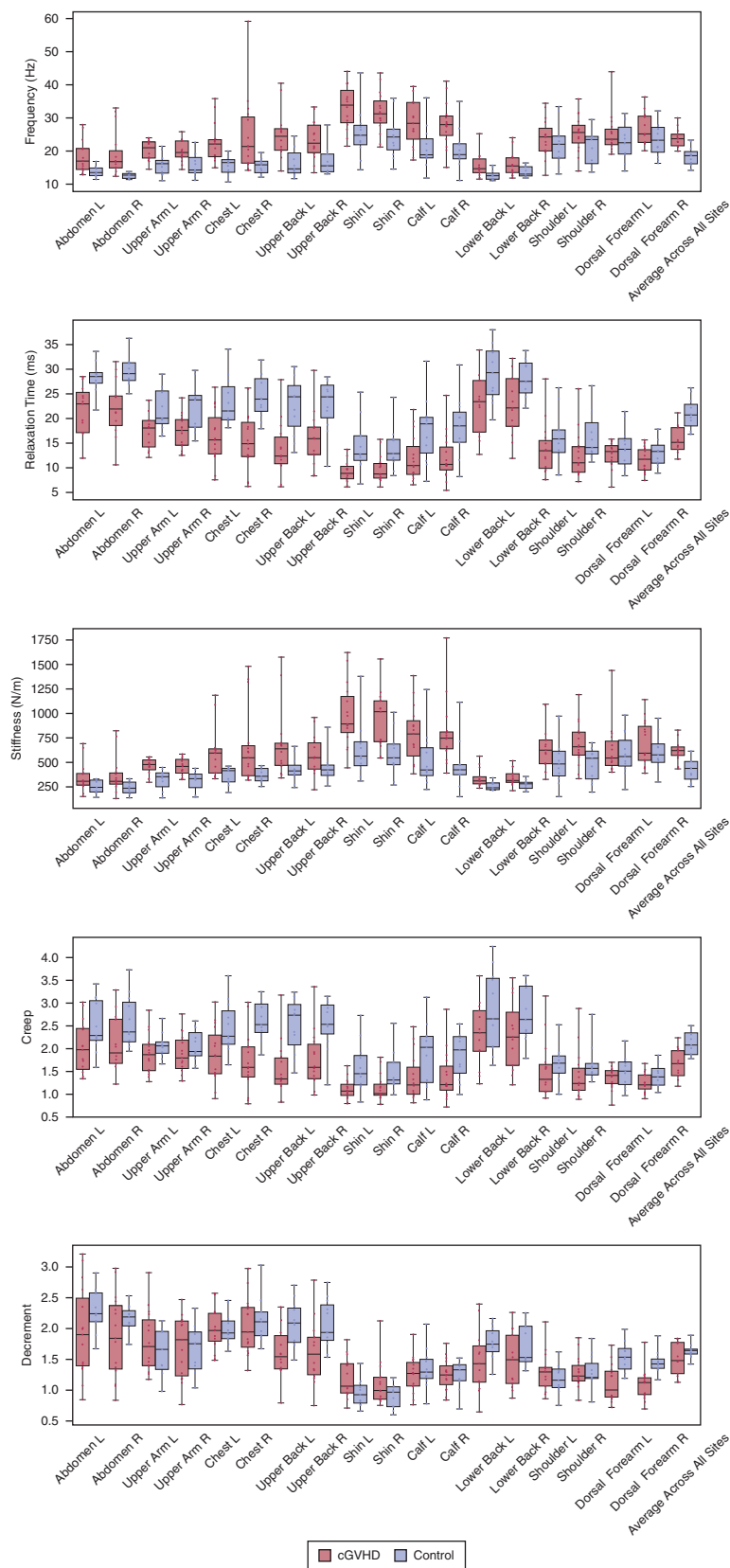
<sup>3</sup>Disease characteristics: AML, ALL, CLL, MDS, MF, AA, and NHL.

<sup>4</sup>Transplant characteristics: PB, BM, and CB.

<sup>5</sup>Days from HCT: days between HCT and MyotonPRO measurement session.

**Figure 1. The 18 measured sites.**  
 Green circles represent the 12 sites that are measured in the supine position.





**Figure 2.** Box and whisker plots of the parameter values by sites in patients with sclerotic cGVHD (red) and in post-HCT controls (blue). Sites are displayed in descending order of the frequency AUCs shown in Figure 4. Bilateral sites are displayed next to each other. The bottom and top boundaries of each box represent the 25th and 75th quartiles, respectively; the horizontal line within the box represents the median; and the whiskers represent the minimum and maximum values. AUC, area under the curve; cGVHD, chronic graft-versus-host disease; HCT, hematopoietic cell transplantation; L, left; R, right.

and relaxation time offers the highest diagnostic yield in differentiating patients with sclerotic cGVHD from post-HCT controls.

Although all models placed importance on frequency and relaxation time, whether these parameters should be used alone or in combination requires further consideration.



**Table 3. Spearman’s Correlation Coefficients between Left and Right Side Parameter Measurements for all the Five Biomechanical Parameters in Patients with Sclerotic cGVHD and in Post-HCT controls**

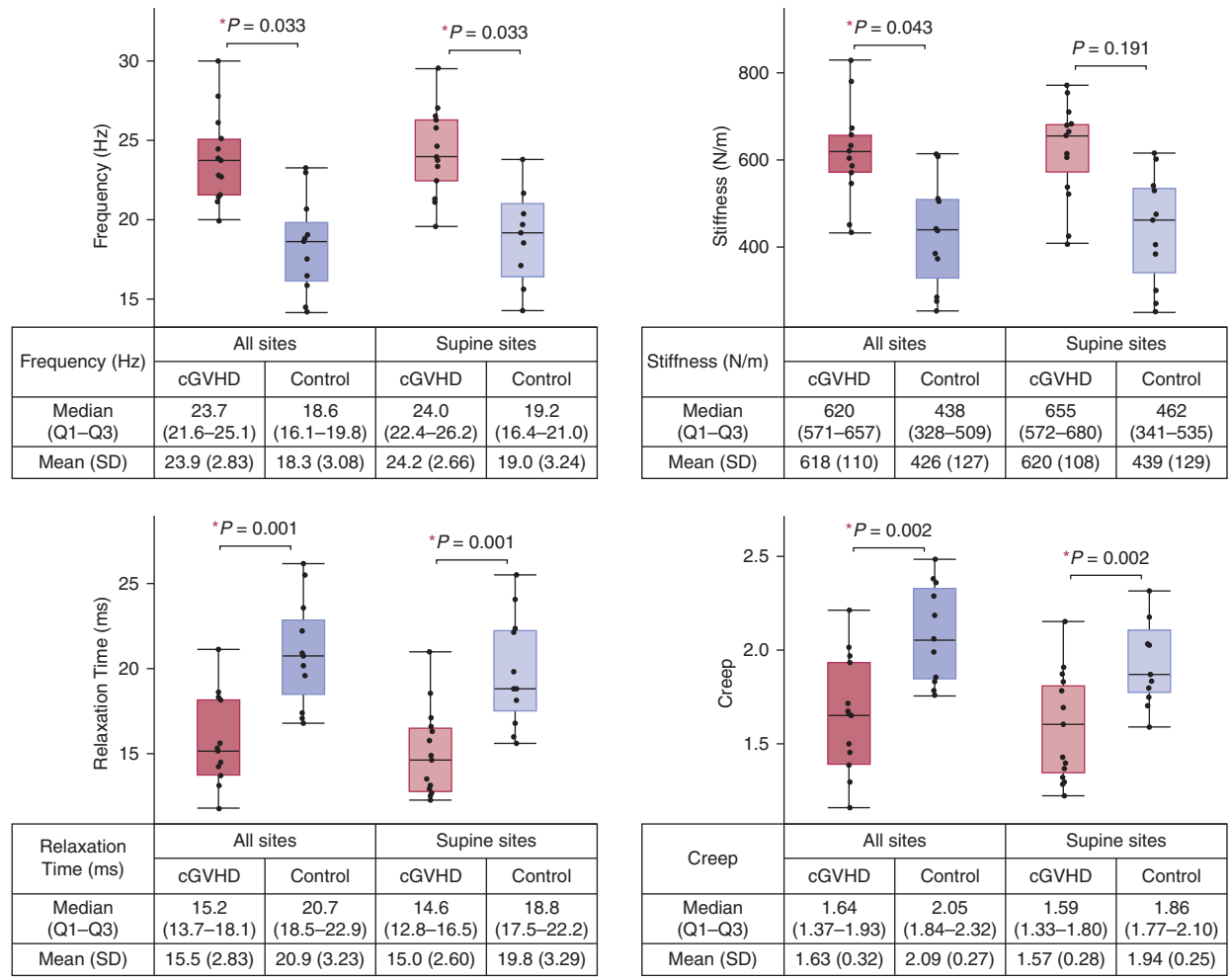
Subject Group	Relaxation Time				
	Frequency	Time	Stiffness	Creep	Decrement
Sclerotic cGVHD	0.812	0.879	0.859	0.847	0.838
Post-HCT controls	0.868	0.915	0.955	0.842	0.843

Abbreviations: cGVHD, chronic graft-versus-host disease; HCT, hematopoietic cell transplantation.

Possible redundant information between frequency and relaxation time, suggested by the high Spearman’s correlation, may limit the added explanatory value provided by the inclusion of the second parameter. Accordingly, with overfit correction, bivariable frequency and relaxation time did not outperform either univariable frequency or univariable

relaxation time analyses. Similarly, backward elimination selected single parameter models (frequency for averages across all the sites and relaxation time for averages across the supine sites) over models with combinations of parameters. In contrast, the LASSO algorithm, which also had the ability to select either a single variable or a combination of variables, favored models using the combination of frequency and relaxation time over single-variable models. An independent data set is needed to validate and compare the classification performance of frequency, relaxation time, and the combination of the two parameters.

This study also informs the development of future myotonometer measurement protocols. Because the AUCs of the parameter averages across the 12 supine sites did not differ significantly from those across all the 18 measured sites, an expedient supine-only measurement protocol may be implemented to reduce measurement time and improve patient experience while maintaining diagnostic accuracy. In addition, the dorsal forearm, which had previously been found to have low interobserver repeatability (Dellalana



**Figure 3.** Distributions of biomechanical parameter values for sclerotic cGVHD patients and post-HCT controls. Patients with sclerotic cGVHD (red boxes) had increased stiffness and frequency and decreased relaxation time and creep compared with post-HCT controls (blue boxes). Patients with sclerotic cGVHD and post-HCT controls were compared by Wilcoxon rank-sum test using averages across the 18 sites (solid boxes, n for patients with cGVHD = 13, n for post-HCT control = 11) and averages across the 12 supine sites (diagonal pattern, n for patients with cGVHD = 15, n for post-HCT control = 11); \*P < 0.05. Frequency and stiffness (top row) values are expected to be higher for sclerotic skin than for healthy skin, whereas relaxation time and creep (bottom row) values are expected to be lower for sclerotic skin. Logarithmic decrement (overall: P = 0.209, supine: P = 0.209) is not shown. cGVHD, chronic graft-versus-host disease; HCT, hematopoietic cell transplantation; Q, quartile.

**Table 4. Univariable and Bivariable Logistic Regression AUCs (with and without Overfit Correction) and 95% CIs for Averages across all the Measured Sites (n for Sclerotic cGVHD = 13, n for Post-HCT controls = 11) and Averages across the 12 Supine Sites (n for Sclerotic cGVHD = 15, n for Post-HCT Controls = 11)**

Parameter(s)	Across all the 18 Measured Sites			Across the 12 Supine Sites		
	AUC (95% CI)	AUC with Overfit Correction	P-Value	AUC (95% CI)	AUC with Overfit Correction	P-Value
Frequency	0.909 (0.786–1.000)	0.910	0.014	0.873 (0.725–1.000)	0.873	0.011
Relaxation Time	0.888 (0.756–1.000)	0.890	0.011	0.873 (0.737–1.000)	0.874	0.011
Stiffness	0.874 (0.734–1.000)	0.875	0.012	0.842 (0.692–0.993)	0.840	0.012
Creep	0.860 (0.711–1.000)	0.862	0.014	0.767 (0.581–0.953)	0.763	0.026
Decrement	0.615 (0.361–0.869)	0.594	0.184	0.576 (0.347–0.804)	0.503	0.497
Frequency and relaxation time	0.916 (0.801–1.000)	0.890	—	0.879 (0.732–1.000)	0.851	—

Abbreviations: AUC, area under the curve; cGVHD, chronic graft-versus-host disease; CI, confidence interval; HCT, hematopoietic cell transplantation.

et al., 2019), exhibited the lowest AUCs for both frequency and relaxation time in this study. Thus, the dorsal forearm may be omitted or substituted with another site, such as the volar forearm, in future measurement protocols. Finally, we report the distribution of values of biomechanical parameters for patients with sclerotic cGVHD and for post-HCT controls, which could inform the design of future myotonometer studies.

This study has several limitations. Of the 11 post-HCT controls, 3 subjects were measured within 12 months of HCT (Table 2) and could have had subclinical cGVHD at the time of measurement. In addition, the myotonometer does not distinguish between cGVHD changes in the skin, changes in subcutaneous tissue, and changes in fascia and skeletal muscle (Hu et al., 2018; Oda et al., 2009). The measurements represent contributions from all underlying

**Table 5. LASSO Regression Coefficients and AUCs of 10 Repeats of 3-Fold Cross-Validation for Averages across All the 18 Measured Sites (n for Sclerotic cGVHD = 13, n for Post-HCT Controls = 11) and Averages across the 12 Supine Sites (n for Sclerotic cGVHD 15, n for Post-HCT Controls 11)**

**LASSO Regression Models for Averages across All the 18 Measured Sites**

Rep	AUC without Overfit Correction (95% CI)	AUC with Overfit Correction (95% CI)	Parameters Selected by Model				
			Frequency	Relaxation Time	Stiffness	Creep	Decrement
1	0.909 (0.786–1)	0.871 (0.641–0.998)	1.722	–0.050	—	—	—
2	0.909 (0.786–1)	0.871 (0.641–0.998)	1.592	–0.030	—	—	—
3	0.909 (0.786–1)	0.871 (0.641–0.998)	1.300	—	—	—	—
4	0.909 (0.786–1)	0.871 (0.641–0.998)	1.456	–0.008	—	—	—
5	0.909 (0.786–1)	0.871 (0.641–0.998)	1.592	–0.030	—	—	—
6	0.909 (0.786–1)	0.871 (0.641–0.998)	1.524	–0.019	—	—	—
7	0.909 (0.786–1)	0.871 (0.641–0.998)	1.904	–0.076	—	—	—
8	0.909 (0.786–1)	0.871 (0.641–0.998)	1.040	—	—	—	—
9	0.909 (0.786–1)	0.871 (0.641–0.998)	1.383	—	—	—	—
10	0.909 (0.786–1)	0.871 (0.641–0.998)	1.592	–0.030	—	—	—

**LASSO Regression Models for Averages across the 12 Supine Sites**

Rep	AUC without Overfit Correction (95% CI)	AUC with Overfit Correction (95% CI)	Parameters Selected by Model				
			Frequency	Relaxation Time	Stiffness	Creep	Decrement
1	0.873 (0.725–1)	0.778 (0.429–0.978)	1.161	–0.139	—	—	—
2	0.873 (0.725–1)	0.778 (0.429–0.978)	0.970	–0.014	—	—	—
3	0.873 (0.725–1)	0.778 (0.429–0.978)	1.040	–0.061	—	—	—
4	0.867 (0.721–1)	0.777 (0.429–0.978)	1.627	–2.367	–0.542	1.500	–0.095
5	0.873 (0.725–1)	0.778 (0.429–0.978)	1.005	–0.038	—	—	—
6	0.873 (0.725–1)	0.778 (0.429–0.978)	1.161	–0.139	—	—	—
7	0.873 (0.725–1)	0.778 (0.429–0.978)	1.040	–0.061	—	—	—
8	0.873 (0.725–1)	0.778 (0.429–0.978)	1.161	–0.139	—	—	—
9	0.873 (0.725–1)	0.778 (0.429–0.978)	0.923	—	—	—	—
10	0.873 (0.725–1)	0.778 (0.429–0.978)	1.040	–0.061	—	—	—

Abbreviations: AUC, area under the curve; cGVHD, chronic graft-versus-host disease; HCT, hematopoietic cell transplantation; LASSO, least absolute shrinkage and selection operator; Rep, repetition.

**Table 6. MyotonPRO Biomechanical Parameters in Descending Order of Feature Importance (1 = most important, 5 = least important) Selected by Random Forest Models (10 iterations) for Averages across All the Sites (n for cGVHD = 13, n for control = 11) and Averages Across the Supine Sites (n for cGVHD = 15, n for control = 11)**

Averages Across All the Sites	
Parameter	Feature Importance
Frequency	1
Relaxation time	2
Stiffness	4
Creep	3
Decrement	5

Averages Across the Supine Sites	
Parameter	Feature Importance
Frequency	1
Relaxation time	2
Stiffness	3
Creep	4
Decrement	5

Abbreviation: cGVHD, chronic graft-versus-host disease. Note the conserved order apart for the position of stiffness and creep.

soft tissue layers, and the proportional contribution of each layer is unknown. Our study design lacked an independent validation data set. Although bootstrapping was used to generate confidence intervals, an independent validation set

is required to assess model accuracy. Finally, the device’s ability to differentiate clinically sclerotic sites from clinically normal sites is an important question precluded by our study design. For clinical implementation, a future study that follows patients before clinical signs of sclerosis through the subsequent appearance of sclerosis is required.

In conclusion, this study shows that a brief clinical protocol measuring the frequency and relaxation time of patients’ skin can effectively differentiate patients with sclerotic cGVHD from post-HCT controls. A larger follow-up study is needed for model validation.

**MATERIALS AND METHODS**

**Data collection**

The study protocol was approved by the Institutional Review Board at Vanderbilt (Institutional Review Board number: 170456) and the U.S. Department of Veterans Affairs Tennessee Valley Healthcare System, Nashville, TN (Institutional Review Board number: 1170055). A total of 15 patients with sclerotic cGVHD (National Institutes of Health skin features scores 2–3) and 11 post-HCT controls were recruited. All participants provided written informed consent. Subjects underwent MyotonPRO measurement by a single observer (F.C.) according to our previously published protocol (Chen et al., 2019a) with a 12 mm disk attachment. Each subject was measured on 9 anatomic sites bilaterally over completely relaxed muscles (shin over tibialis anterior, dorsal forearm over extensor digitorum, upper arm over biceps brachii caput longum, shoulder over medial deltoideus, chest over pectoralis major, abdomen over rectus abdominus, calf over gastrocnemius caput laterale, upper back over medial trapezius, and lower back over erector spinae), resulting in 18 total measurement sites. Each measurement session

**Table 7. Representative Combinations for Parameters Selected by LASSO Regression as the Most Important Variables in Differentiating Patients with Sclerotic cGVHD from Post-HCT Controls for Each Site**

Sites	Frequency	Relaxation Time	Stiffness	Creep	Decrement	AUC with Overfit Correction (95% CI)
Abdomen L	+	+	+		+	0.884 (0.674–0.996)
Abdomen R	+	+			+	0.902 (0.737–0.999)
Upper Arm L	+		+			0.773 (0.431–0.981)
Upper Arm R	+		+			0.773 (0.422–0.983)
Chest L	+	+		+	+	0.801 (0.437–0.992)
Chest R		+				0.829 (0.546–0.992)
Upper back L	+				+	0.840 (0.533–0.994)
Upper back R	+				+	0.756 (0.358–0.987)
Shin L		+			+	0.855 (0.606–0.992)
Shin R		+	+		+	0.761 (0.407–0.978)
Calf L	+					0.741 (0.404–0.964)
Calf R	+	+	+	+	+	0.780 (0.411–0.986)
Lower back L		+	+		+	0.812 (0.445–0.988)
Lower back R	+		+	+	+	0.740 (0.327–0.979)
Shoulder L	+	+	+	+	+	0.725 (0.381–0.955)
Shoulder R	+	+	+	+	+	0.757 (0.392–0.981)
Dorsal forearm L		+			+	0.782 (0.427–0.985)
Dorsal Forearm R	+				+	0.846 (0.559–0.995)
Average across all sites	+	+				0.871 (0.641–0.998)
Average across supine sites	+	+				0.778 (0.429–0.978)

Abbreviations: AUC, area under the curve; cGVHD, chronic graft-versus-host disease; CI, confidence interval; HCT, hematopoietic cell transplantation; L, left; LASSO, least absolute shrinkage and selection operator; R, right. Representative: of 10 repetitions, the most commonly selected combination of parameters for each site is shown.



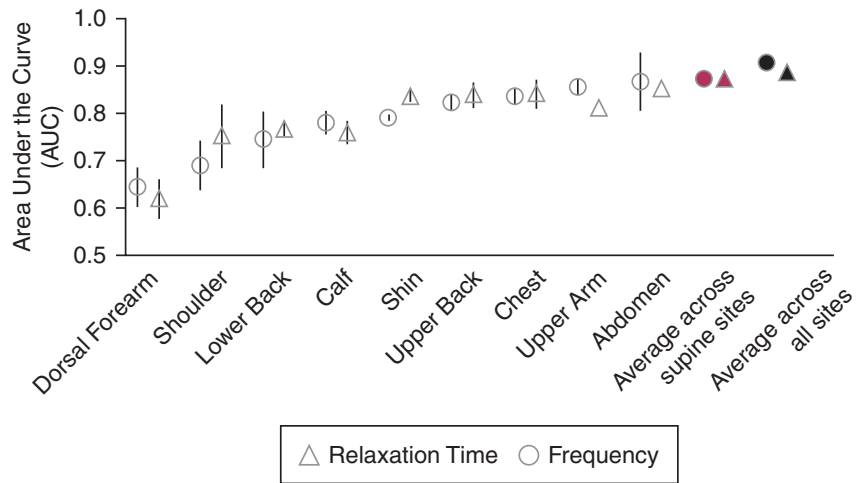
**Table 8. AUC of Univariable Analyses of Each of the Five Parameters for Individual Sites**

Site	Frequency		Relaxation Time		Stiffness		Creep		Decrement	
	AUC (95% CI)	Overfit-Corrected AUC	AUC (95% CI)	Overfit-Corrected AUC	AUC (95% CI)	Overfit-Corrected AUC	AUC (95% CI)	Overfit-Corrected AUC	AUC (95% CI)	Overfit-Corrected AUC
Abdomen L	0.824 (0.665–0.983)	0.825	0.873 (0.737–1)	0.871	0.661 (0.437–0.884)	0.652	0.739 (0.543–0.936)	0.728	0.661 (0.436–0.885)	0.628
Abdomen R	0.939 (0.845–1)	0.949	0.873 (0.725–1)	0.878	0.745 (0.550–0.941)	0.743	0.758 (0.567–0.949)	0.753	0.606 (0.376–0.836)	0.572
Upper arm L	0.861 (0.715–1)	0.852	0.788 (0.610–0.966)	0.778	0.855 (0.702–1)	0.848	0.667 (0.452–0.882)	0.627	0.600 (0.373–0.827)	0.543
Upper arm R	0.824 (0.656–0.992)	0.818	0.776 (0.593–0.959)	0.766	0.842 (0.693–0.992)	0.833	0.648 (0.431–0.866)	0.607	0.582 (0.350–0.814)	0.493
Chest L	0.861 (0.718–1)	0.853	0.836 (0.675–0.998)	0.828	0.812 (0.642–0.983)	0.801	0.779 (0.599–0.958)	0.769	0.491 (0.260–0.722)	0.398
Chest R	0.812 (0.645–0.979)	0.810	0.885 (0.755–1)	0.884	0.812 (0.646–0.978)	0.804	0.873 (0.730–1)	0.874	0.588 (0.359–0.817)	0.51
Upper back L	0.851 (0.697–1)	0.852	0.877 (0.739–1)	0.877	0.779 (0.594–0.965)	0.777	0.877 (0.731–1)	0.877	0.799 (0.622–0.976)	0.795
Upper back R	0.821 (0.648–0.995)	0.819	0.818 (0.624–1)	0.815	0.727 (0.515–0.939)	0.706	0.825 (0.632–1)	0.824	0.773 (0.583–0.963)	0.768
Shin L	0.8 (0.609–0.991)	0.788	0.836 (0.662–1)	0.829	0.824 (0.643–1)	0.815	0.818 (0.639–0.997)	0.811	0.733 (0.533–0.933)	0.718
Shin R	0.788 (0.595–0.98)	0.779	0.824 (0.661–0.987)	0.816	0.861 (0.719–1)	0.853	0.770 (0.584–0.956)	0.762	0.664 (0.443–0.884)	0.639
Calf L	0.782 (0.593–0.97)	0.771	0.764 (0.563–0.964)	0.747	0.782 (0.583–0.980)	0.762	0.721 (0.510–0.932)	0.7	0.582 (0.349–0.814)	0.514
Calf R	0.818 (0.633–1)	0.808	0.800 (0.611–0.989)	0.788	0.861 (0.690–1)	0.857	0.752 (0.553–0.95)	0.738	0.491 (0.251–0.731)	0.399
Lower back L	0.804 (0.623–0.986)	0.801	0.776 (0.586–0.966)	0.773	0.790 (0.605–0.975)	0.786	0.657 (0.427–0.888)	0.628	0.713 (0.492–0.935)	0.692
Lower back R	0.685 (0.465–0.906)	0.665	0.755 (0.556–0.955)	0.75	0.783 (0.593–0.973)	0.781	0.699 (0.484–0.914)	0.671	0.678 (0.453–0.904)	0.649
Shoulder L	0.594 (0.360–0.828)	0.517	0.655 (0.428–0.881)	0.607	0.673 (0.449–0.896)	0.627	0.630 (0.398–0.862)	0.554	0.567 (0.334–0.799)	0.504
Shoulder R	0.691 (0.478–0.904)	0.681	0.752 (0.561–0.942)	0.741	0.776 (0.595–0.956)	0.779	0.733 (0.528–0.938)	0.706	0.515 (0.279–0.751)	0.422
Dorsal forearm L	0.594 (0.350–0.838)	0.545	0.570 (0.323–0.817)	0.515	0.503 (0.260–0.746)	0.433	0.618 (0.370–0.866)	0.568	0.824 (0.663–0.985)	0.832
Dorsal forearm R	0.685 (0.465–0.905)	0.671	0.648 (0.427–0.870)	0.618	0.594 (0.367–0.821)	0.555	0.673 (0.458–0.887)	0.650	0.915 (0.788–1)	0.918
Average across 18 sites	<b>0.909 (0.786–1)</b>	<b>0.910</b>	<b>0.888 (0.756–1)</b>	<b>0.890</b>	<b>0.874 (0.734–1)</b>	<b>0.875</b>	<b>0.860 (0.711–1)</b>	<b>0.862</b>	<b>0.615 (0.361–0.869)</b>	<b>0.594</b>

Abbreviation: AUC, area under the curve; CI, confidence interval; L, left; R, right.

**Figure 4. AUC of univariable analyses of frequency and relaxation time for individual anatomic sites.**

Frequency = open circles, and relaxation time = open triangles. Averages across the 12 supine sites (red-filled markers, n for patients with cGVHD = 15; n for controls = 11) and across all the 18 measured sites (black-filled markers, n for patients with cGVHD = 13, n for controls = 11) are shown. Markers represent the averages of the left and right anatomic sites' AUCs. Error bars represent the ranges between the left and right AUCs. AUC, area under the curve.



lasted ~20 minutes. Two of the patients with cGVHD were measured only on the 12 sites that could be measured with the patient in the supine position (bilateral shin, dorsal forearm, upper arm, shoulder, chest, and abdomen) owing to discomfort when laying in the prone position. These supine-only measurement sessions lasted ~10–15 minutes.

**Statistical analysis**

For each subject, we calculated the parameter averages over all the 18 measured sites and over the 12 supine sites. The Wilcoxon rank-sum test was performed to compare the distributions of parameter averages between patients with sclerotic cGVHD and post-HCT controls. Spearman's correlation coefficient was calculated for each pair of the parameters. We evaluated the discriminatory ability of the parameters using (i) univariable logistic regression for each parameter, (ii) backward stepwise selection on the five-parameter logistic regression model, (iii) bivariable logistic

regression consisting of frequency and relaxation time, (iv) LASSO regression, and (v) random forest. The AUC was used to compare model performances. Overfit-corrected AUCs were calculated from bootstrap 0.632+ method with 100 bootstrap replications (Efron and Tibshirani, 1997). All analyses were completed using R, version 3.6.1.

A backward stepwise selection was performed by eliminating one parameter at a time from the full five-parameter model to minimize the AIC. The AIC assesses the models' goodness of fit while penalizing model complexity. A smaller AIC value suggests a simpler model with a better fit. The final model is reached when the elimination of any remaining parameter no longer reduces the AIC.

LASSO is a machine learning algorithm that extends standard logistic regression models by enabling variable selection in the fitting process (Friedman et al., 2010). All variables were standardized before LASSO regression. The R package glmnet was applied to build a LASSO regression (Friedman et al., 2021). Each analysis

**Table 9. Spearman's Correlation Coefficient Matrix between the MyotonPRO Biomechanical Parameters Using Parameter Averages in all Study Patients (Combined Sclerotic cGVHD and Post-Transplant Control) for Averages across All Sites (n = 24) and Averages Across Supine Sites (n = 26)**

Averages Across All Sites					
	Frequency	Relaxation Time	Stiffness	Creep	Decrement
Frequency	1	-0.94	0.95	-0.86	-0.28
Relaxation time	-0.94	1	-0.93	0.94	0.41
Stiffness	0.95	-0.93	1	-0.82	-0.22
Creep	-0.86	0.94	-0.82	1	0.54
Decrement	-0.28	0.41	-0.22	0.54	1
Averages Across Supine Sites					
	Frequency	Relaxation Time	Stiffness	Creep	Decrement
Frequency	1	-0.94	0.95	-0.87	-0.10
Relaxation time	-0.94	1	-0.94	0.94	0.15
Stiffness	0.95	-0.94	1	-0.87	-0.04
Creep	-0.87	0.94	-0.87	1	0.32
Decrement	-0.10	0.15	-0.04	0.32	1

Abbreviation: cGVHD, chronic graft-versus-host disease.

Higher absolute values of Spearman's correlation coefficient represent a greater correlation between the two parameters.

**Table 10. Topics of Investigation Required for the Myotonometer to Serve as a Valid Response Measure Instrument in the Management of Sclerotic cGVHD**

Serial Number	Topics of Investigation	Relevant Studies
1	Intraoperator and interoperator reliability	Dellalana et al., 2019.
2	Ability to discriminate obvious patients with sclerotic from healthy patients	Chen et al., 2019a.
3	Ability to discriminate patients with obvious sclerosis from post-HCT controls without clinical evidence of sclerosis	The current study
4	Determination of the optimal parameter or combination of parameters for topic of investigation 3 (above)	The current study
5	Selection of an efficient clinical protocol that future studies and clinical practice can use for patient-level investigation	Chen et al., 2019a. The current study found that a supine-only protocol maintained discriminatory ability.
6	Determination of whether individual sites are more or less useful than averages across sites. Determination of how individual sites correlate with the clinical appreciation of sclerosis at or near those sites	The current study provides the typical ranges of parameter values that may be used to design future studies.
7	Consistency of longitudinal measurements with clinical response	Baker et al., 2021
8	Ability to detect subclinical disease	Future study

Abbreviations: cGVHD, chronic graft-versus-host disease; HCT, hematopoietic cell transplantation.

consisted of 10 repetitions of three-fold cross-validation. The three-fold cross-validation selected the tuning parameters and determined which of the five biomechanical properties, alone or in combination, best indicated the diagnosis of sclerotic cGVHD.

Random forest is an ensemble learning algorithm that uses sets of classification or regression trees to evaluate the potential variables and rank them according to their relative strength in predicting the outcome (Strobl et al., 2009). The R package randomForest was applied to build the random forest classification algorithm (Breiman et al., 2018). Random forest analysis ranked the five biomechanical parameters by their importance in differentiating patients with sclerotic cGVHD from post-HCT controls.

#### Data availability statement

Datasets related to this article can be found at <https://doi.org/10.17632/6mdy96dkxz.1>, an open-source online data repository hosted at Mendeley Data (Baker et al., 2021).

#### ORCIDiS

Laura X. Baker: <http://orcid.org/0000-0002-0655-0164>  
 Fuyao Chen: <http://orcid.org/0000-0002-9609-5731>  
 Austin Cronin: <http://orcid.org/0000-0002-6855-4582>  
 Heidi Chen: <http://orcid.org/0000-0001-5668-2758>  
 Arved Vain: <http://orcid.org/0000-0002-8031-424X>  
 Madan Jagasia: <http://orcid.org/0000-0001-5757-5604>  
 Eric R. Tkaczyk: <http://orcid.org/0000-0002-2850-4740>

#### AUTHOR CONTRIBUTIONS

Conceptualization: ERT, MJ, FC; Data Curation: FC; Formal analysis: HC, LXB; Funding Acquisition: ERT; Methodology: ERT, AV; Supervision: ERT; Validation: AC; Visualization: LXB; Writing - Original Draft Preparation: LXB; Writing - Review and Editing: LXB, FC, AC, HC, AV, MJ, ERT

#### ACKNOWLEDGMENTS

This work was supported by Career Development Award number IK2 CX001785 from the U.S. Department of Veterans Affairs Clinical Science research and Development Service and National Institutes of Health K12 CA090625. The authors are grateful to the patients and volunteers who participated in this study.

#### CONFLICT OF INTEREST

AV is the inventor of the Myoton but has no financial interest in the device. The remaining authors state no conflict of interest.

#### REFERENCES

Arai S, Arora M, Wang T, Spellman SR, He W, Couriel DR, et al. Increasing incidence of chronic graft-versus-host disease in allogeneic transplantation:

a report from the Center for International Blood and Marrow Transplant Research. *Biol Blood Marrow Transplant* 2015;21:266–74.

Baker LX, Chen F, Ssempijja Y, Byrne M, Kim TK, Vain A, et al. Longitudinal tracking of skin dynamic stiffness to quantify evolution of sclerosis in chronic graft-versus-host disease. *Bone Marrow Transplant* 2021;56:989–91.

Breiman L, Cutler A, Liaw A, Wiener M. randomForest: Breiman and Cutler's random forests for classification and regression. <https://cran.r-project.org/web/packages/randomForest/index.html>; 2018 (accessed 6 July, 2020).

Carpenter PA, Kitko CL, Elad S, Flowers ME, Gea-Banacloche JC, Halter JP, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: V. The 2014 ancillary therapy and supportive care working group report. *Biol Blood Marrow Transplant* 2015;21:1167–87.

Chen F, Dellalana LE, Gandelman JS, Vain A, Jagasia MH, Tkaczyk ER. Non-invasive measurement of sclerosis in cutaneous cGVHD patients with the handheld device Myoton: a cross-sectional study [published correction appears in *Bone Marrow Transplant* 2020;55:992]. *Bone Marrow Transplant* 2019a;54:616–9.

Chen F, Wang L, Vain A, Ssempijja Y, Dellalana L, Zhang K, et al. Interobserver reproducibility of the Myoton and Durometer devices to measure skin stiffness and hardness in chronic cutaneous graft-versus-host disease patients. *Blood* 2019b;134(Suppl. 1):4515.

Dellalana LE, Chen F, Vain A, Gandelman JS, Pöldemaa M, Chen H, et al. Reproducibility of the durometer and myoton devices for skin stiffness measurement in healthy subjects. *Skin Res Technol* 2019;25:289–93.

Efron B, Tibshirani R. Improvements on cross-validation: the 632+ bootstrap method. *J Am Stat Assoc* 1997;92:548–60.

Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010;33:1–22.

Friedman J, Hastie T, Tibshirani R, Narasimhan B, Tay K, Simon N, et al. glmnet: lasso and elastic-net regularized generalized linear models. <https://cran.r-project.org/web/packages/glmnet/index.html>; 2021 (accessed 6 July, 2020).

Hu X, Lei D, Li L, Leng Y, Yu Q, Wei X, et al. Quantifying paraspinal muscle tone and stiffness in young adults with chronic low back pain: a reliability study. *Sci Rep* 2018;8:14343.

Inamoto Y, Storer BE, Petersdorf EW, Nelson JL, Lee SJ, Carpenter PA, et al. Incidence, risk factors, and outcomes of sclerosis in patients with chronic graft-versus-host disease. *Blood* 2013;121:5098–103.

Oda K, Nakaseko C, Ozawa S, Nishimura M, Saito Y, Yoshida F, et al. Fasciitis and myositis: an analysis of muscle-related complications caused by chronic GVHD after allo-SCT. *Bone Marrow Transplant* 2009;43:159–67.

Odell ID, Flavell RA, Silver FH. Vibrational optical coherence tomography detects unique skin fibrotic states: preliminary results of animal and human

studies [e-pub ahead of print]. *J Am Acad Dermatol* 2020. <https://doi.org/10.1016/j.jaad.2020.08.091> (accessed October 6, 2020).

Socié G, Ritz J. Current issues in chronic graft-versus-host disease. *Blood* 2014;124:374–84.

Socié G, Stone JV, Wingard JR, Weisdorf D, Henslee-Downey PJ, Bredeson C, et al. Long-term survival and late deaths after allogeneic bone marrow transplantation. Late Effects Working Committee of the International Bone Marrow Transplant Registry. *N Engl J Med* 1999;341:14–21.

Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods* 2009;14:323–48.

Vain A, Myoton AS. WO2012/089221: device and method for real-time measurement of parameters of mechanical stress state and biomechanical properties of soft biological tissue. <https://patentscope.wipo.int/search/pt/detail.jsf;jsessionid=7143D7FEDDB5DA78F04E7D2C59509229.wapp1nB?docId=WO2012089221&recNum=1&tab=PCTClaims&maxRec=1&office=&prevFilter=&sortOption=&queryString=WO2012%2F089221;2012> (accessed 1 October, 2020).



**This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>**