Biol Blood Marrow Transplant 26 (2020) 562–567

Biology of Blood and Marrow Transplantation

journal homepage: www.bbmt.org

Survivorship

Reliability and Validity of the Modified 7-Day Lee Chronic Graft-versus-Host Disease Symptom Scale

Christopher Teh, Lynn Onstad, Stephanie J. Lee*

Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington

Article history:
Received 6 September 2019
Accepted 13 November 2019

Keywords:
chronic graft-versus-host disease
symptoms
quality of life
patient-reported outcomes
survey

A B S T R A C T

Chronic graft-versus-host disease (cGVHD) adversely affects patient quality of life, functional status, and survival after allogenic hematopoietic cell transplantation. The Lee Symptom Scale is a 30-item scale developed to measure the symptoms of cGVHD. Although the original 30-item scale uses a 1-month recall period, we tested the reliability and validity of a 28-item scale (deleting 2 items based on supportive care needs rather than symptoms) with a 7-day recall period, a format that is more appropriate for use in clinical trials. Results show the modified 7-day scale is reliable and valid in the modern era and may be used to assess the symptom burden of cGVHD in clinical trials. Using the distribution method, a 5- to 6-point difference (half a standard deviation) is considered clinically meaningful.

INTRODUCTION

Chronic graft-versus-host disease (cGVHD) is a serious iatrogenic complication that affects survivors of allogeneic hematopoietic cell transplant (HCT). Twenty percent to 50% of allogeneic transplant survivors develop cGVHD. The disease results in inflammation, scarring, and organ dysfunction. cGVHD is the most common long-term complication of HCT and is associated with a decreased quality of life, impaired functional status, continued need for immunosuppressive medications, and increased nonrelapse mortality [1].

Published in 2002, the Lee cGVHD Symptom Scale (LSS) was developed to measure symptoms in adult outpatients with cGVHD [2]. The scale contains 30 items grouped into 7 subscales (skin, eye, mouth, lung, nutrition, energy, and psychological) and takes 2 minutes to complete. Patients report how "bothered" they feel about each symptom over the previous month using a 5-point Likert scale from "not at all" to "extremely." A 1-month assessment period was chosen for the original scale to capture symptoms over a period of time because most patients with cGVHD are treated as outpatients, and cGVHD symptoms can wax and wane. Subscales range from 0 to 100, with a higher score indicating worse symptoms. Subscales may be scored if at least 50% of items are answered, and subscales are averaged to calculate the summary score. Readers are cautioned to use the correct scoring algorithm (Table 1) because the headers in the survey do not directly correspond to the subscales. In the original publication a 6- to 7-point change in the summary score suggested a clinically meaningful difference in patient symptomatology.

In 2005 and 2014, the National Institutes of Health (NIH) Consensus Development Project on Criteria for Clinical Trials in cGVHD proposed the LSS as a tool to determine the efficacy of cGVHD treatments [3,4]. The relevance of the scale to modern cGVHD patients was confirmed in a 2016 publication about the content validity of the scale [5]. Even though the LSS is now commonly used to evaluate symptoms in cGVHD prevention [6,7] and therapy trials [8-10], the US Food and Drug Administration and pharmaceutical sponsors prefer a shorter recall period of 7 days. Some have also questioned the inclusion of 2 items (ie, "Receiving nutrition from an intravenous line or feeding tube" and "Need to use oxygen") because they reflect use of supportive care measures rather than symptoms. In general, recall is better over shorter periods that are preferred for symptom assessment but may be influenced by fluctuating symptomatology. A 7-day recall period matches the common quality of life instruments.

The aim of the present study was to reassess the instrument's reliability and validity in the modern era with a 7-day recall period to establish internal consistency of items, show that scores are stable if a patient's condition does not change, and demonstrate convergent and divergent validity. These are important features of any scale used to document effectiveness of treatments. We also evaluated the impact of deleting the 2 questions relating to supportive care on the performance of the scale.

Financial disclosure: See Acknowledgments on page 567.

*Correspondence and reprint requests: Stephanie J. Lee, MD, MPH, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue N, D5-290, Seattle, WA 98109.

E-mail address: sjlee@fredhutch.org (S.J. Lee).

https://doi.org/10.1016/j.bbmt.2019.11.020
1083-8791/

Table 1
Scoring Algorithm for the mLSS

Subscale Name	Number of Items	Items
Skin	5	a. Abnormal skin color b. Rashes c. Thickened skin d. Sores on skin e. Itchy skin
Eye	3	f. Dry eyes g. Need to use eye drops frequently h. Difficulty seeing clearly
Mouth	2	i. Need to avoid certain foods due to mouth pain j. Ulcers in mouth
Lung	4	l. Frequent cough m. Colored sputum o. Shortness of breath at rest ~~p. Need to use oxygen~~ **aa. Fevers**
Nutrition	4	~~**k. Receiving nutrition from an intravenous line or feeding tube**~~ q. Difficulty swallowing solid foods r. Difficulty swallowing liquids s. Vomiting t. Weight loss
Energy	7	**n. Shortness of breath with exercise** **u. Joint and muscle aches** **v. Limited joint movement** **w. Muscle cramps** **x. Weak muscles** y. Loss of energy z. Need to sleep more/take naps
Psych	3	bb. Depression cc. Anxiety dd. Difficulty sleeping

The 7-day mLSS is a 28-item instrument with 7 subscales (skin, eyes, mouth, lung, nutrition, energy, and psych) containing 2 to 7 items that allow calculation of a summary score. Response options for "Please let us know if you have been bothered by any of the following problems in the past 7 days" range from 0 to 4 (Not at all, Slightly, Moderately, Quite a bit, Extremely). A clinically meaningful difference is considered 5-6 points on the summary score. **Bolded** items are scored under a different subscale than where they are located under the headers in the survey. Items p and k were deleted from the original 30 item scale.
Scoring rules:
a. Note that the subscales do not conform exactly to the headers in the patient survey.
b. Items p and k are deleted from the 7-day version.
c. Subscales may be scored if 50% of more of the items in the subscale are completed.
d: Scores are linearly transformed to a 0-100 scale where 0 means all answered items were a "0" and "100" means that all answered items were a "4."
e. Missing items are not included in the scoring.
f. The summary score is the average of the subscale scores, as long as 4 or more subscales are available.
g. Higher scores indicate more severe symptoms.

METHODS

Sixty-eight participants were included in the study if they were adults aged ≥ 18 years, able to communicate in English, diagnosed with cGVHD per the 2014 NIH consensus criteria, and symptomatic with active cGVHD with an NIH score > 1 in one or more organs associated with cGVHD. We anticipated enrolling up to 80 subjects to include at least 40 of those with unchanged cGVHD symptoms between the first and second administration of the LSS. Data collection continued until 40 participants indicated no clinically significant change to their symptoms ("about the same") between their baseline and follow-up surveys. The study was approved by the Fred Hutchinson Institutional Review Board and was conducted between 2016 and 2017. The requirement for written documentation of consent was waived given the minimal risk nature of the study because participants were informed of all components of informed consent, including that they could skip over any topic they wished and that participation was voluntary and would not affect their care.

The LSS survey used in this study was identical to the published 1-month recall period version except the recall period used was "the past 7 days." Participants were given a paper survey at enrollment that they completed in the clinic or mailed back. They then were mailed a second survey to complete and return by mail approximately 1 week after completion of the first LSS. The follow-up survey asked whether their cGVHD symptoms had changed since enrollment, using a 7-point scale: very much worse, moderately worse, a little worse, about the same, a little better, moderately better, or very much better. Patients also completed 1 page of sociodemographic questions with their enrollment survey. On both surveys they answered questions about their overall cGVHD status from the NIH patient self-report cGVHD assessment that included reporting whether they considered their cGVHD "mild," "moderate," or "severe."

Descriptive statistics include patient, transplant, and cGVHD characteristics. The survey was scored according to the published recommendations and excluding the 2 supportive care items in question [2,5]; nonresponse was defined as the inability to calculate a score because of missing data. The test-retest correlation was calculated for the summary score and the 7 subscales and reported as the Spearman correlation coefficient. The intraclass correlations are reported as Cronbach's alphas. Generally, test-retest and Cronbach's alpha values of >.7 [11] are considered acceptable.

RESULTS

During the study 68 patients enrolled, and 40 (59%) reported that their symptoms were "about the same" on the follow-up survey. The other 28 patients either failed to complete a follow-up survey (n = 12, 18%) or indicated on the follow-up survey that their cGVHD symptoms had improved (n = 16, 24%). No one reported that their symptoms had worsened. Psychometrics are reported based on the 68 enrollment surveys, whereas the test-retest statistics are based on the 40 participants who reported that their cGVHD had not changed since they first completed the LSS. Sample characteristics are shown in Table 2. Ten participants (14.7%) identified themselves as nonwhite: Asian (n = 3), black (n = 3), Hawaiian Native/Pacific Islander (n = 1), and other (n = 3). Forty-three participants (64.2%) had a college or postgraduate degree, and 43 (64.2%) were married or living with a partner. Twenty-two participants (32.8%) were working full-time and 6 (9.0%) part-time.

The median time elapsed from HCT to the diagnosis of cGVHD was 8.4 months. The median time from HCT to enrollment was 34.9 months (interquartile range, 19.4 to 64), and the distribution of global severity of cGVHD using 2014 NIH consensus scoring was mild (n = 10), moderate (n = 30), and severe (n = 28). NIH severity was higher than participant self-reported severity, which was none (n = 5), mild (n = 31), moderate (n = 29), and severe (n = 2). The most common organs with scores of 2 or higher were skin in 31 participants (45.6%) and eye in 23 (33.8%).

Table 3 summarizes the psychometric properties of the survey for both the 30-item and 28-item scales. Cronbach's alpha was >.7 for the energy, skin, eye, and mouth subscores and for the summary score but <.62 for nutrition, lung, and psychological scales. No participants endorsed the intravenous or feeding tube item, and 2 reported being "slightly" or "moderately" bothered by needing to use oxygen. Removing these items and recalculating the subscale scores minimally improved the Cronbach's alpha of the nutrition and lung subscales to .61 and .43, respectively when compared with inclusion of all items. Importantly, however, Cronbach's alpha of the summary score remained high and was .84 for the 30-item scale and .85 for the 28-item scale in the present study compared to .90 in the original description. Cronbach's alpha remained .76 to .83 when evaluated in the 3 severity groups separately. The standard deviation was 10.5 for the 30-item scale and 10.7 for the 28-item scale; we estimate that a 5- to 6-point difference is clinically meaningful using the distribution method (half a standard deviation) [12,13].

Table 2
Cohort Characteristics (N = 68)

Characteristic	All Participants (N = 68)	Participants Used for Test-Retest (n = 40)
Median age, yr (IQR)	57.5 (42.5-63.5)	58.5 (40.5-63)
Male sex	41 (60.3)	26 (65.0)
Race/ethnicity		
White	58 (85.3)	37 (92.5)
Asian	3 (4.4)	1 (2.5)
Black	3 (4.4)	1 (2.5)
Hawaiian Native/Pacific Islander	1 (1.5)	0 (0)
Other	3 (4.4)*	1 (2.5)†
Hispanic	0 (0)	0 (0)
Marital status		
Married/living with partner	43 (64.2)	27 (67.5)
Single, never married	14 (20.9)	7 (17.5)
Divorced, separated	8 (11.9)	6 (15.0)
Widowed	1 (1.5)	0 (0)
Married/not living with partner	1 (1.5)	0 (0)
Missing	(n = 1)	
Education		
Less than college	8 (11.9)	5 (13)
Some college	16 (23.9)	9 (22.5)
College graduate	24 (35.8)	14 (35.0)
Post graduate degree	19 (28.4)	12 (30.0)
Missing	(n = 1)	
Work/school status		
Working or school full time	22 (32.8)	15 (37.5)
Working part time	6 (9.0)	3 (7.5)
Retired	16 (23.9)	10 (25.0)
Disabled, unable to work	12 (17.9)	8 (20.0)
Homemaker	7 (10.4)	3 (7.5)
On medical leave	2 (3.0)	0 (0)
Unemployed, looking for work	2 (3.0)	1 (2.5)
Missing	(n = 1)	(n = 0)
NIH severity at enrollment (patient self-report)		
Mild	36 (53.7)	21 (52.5)
Moderate	29 (43.3)	17 (42.5)
Severe	2 (3.0)	2 (5.0)
Missing	(n = 1)	
NIH severity at enrollment (per NIH criteria)		
Mild	10 (14.7)	5 (12.5)
Moderate	30 (44.1)	16 (40.0)
Severe	28 (41.2)	19 (47.5)
Median time from transplant to cGVHD, mo (IQR)	8.4 (5.4-11.8)	10.0 (5.8-12.3)
Median time from cGVHD diagnosis to enrollment, mo (IQR)	34.9 (19.4-64)	34.9 (19.6-60.5)
Score 2-3 organ involvement		
Skin	31 (45.6)	19 (47.5)
Eye	23 (33.8)	15 (37.5)
Mouth	5 (7.4)	2 (5.0)
Gastrointestinal	1 (1.5)	1 (2.5)
Liver (1 missing)	0 (0)	0 (0)
Lung (1 missing)	7 (10.4)	6 (15.4)
Joint	12 (17.6)	8 (20.0)

Values are n (%) unless otherwise defined. IQR indicates interquartile range.
* Asian and Indian, Portuguese, and American.
† American.

Table 3
Reliability of the 7-Day LSS (N = 68)

			Nutrition		Lung					Summary	
	Energy	Skin	Original	Modified	Original	Modified	Psych	Eye	Mouth	Original	Modified
Items	7	5	5	4	5	4	3	3	2	30	28
Mean	28.0	15.0	5.0	6.3	3.4	3.9	16.5	44.2	19.9	18.9	19.1
Standard deviation	20.2	17.0	8.5	10.6	6.2	7.4	16.6	28.6	25.0	10.5	10.7
Median	25.0	10.0	0	0	0	0	12.5	41.7	12.5	17.6	18.0
Range	0-85.7	0-70	0-45	0-56.3	0-40	0-50	0-75	0-100	0-100	2.4-43.3	2.4-43.5
Cronbach's α	.85	.74	.57	.61	.40	.43	.57	.83	.71	.84	.85
Floor, n (%)	5 (7.4%)	21 (30.9%)	40 (58.8%)	40 (58.8%)	41 (60.3%)	42 (61.8%)	18 (26.5%)	7 (10.3%)	29 (42.6%)	1 (1.5%)	1 (1.5%)
Ceiling, n (%)	1 (1.5%)	1 (1.5%)	1 (1.5%)	1 (1.5%)	1 (1.5%)	1 (1.5%)	1 (1.5%)	4 (5.9%)	1 (1.5%)	1 (1.5%)	1 (1.5%)
Nonresponse	0	0	0	0	0	0	0	0	0	0	0
Test-retest (n = 40)	.89	.88	.76	.76	.71	.70	.82	.85	.79	.79	.81
Intercorrelation											
Energy		.38*	.50†	.50†	.38*	.37*	.35*	.27	.22	.71†	.71†
Skin			.01	.01	.11	.13	.27	.07	.14	.46†	.46†
Nutrition					.24		.21	.19	.46†	.56†	
Nutrition-modified						.28	.21	.19	.46†		.57†
Lung							.01	.10	−.03	.29	
Lung-modified							.05	.05	.01		.29
Psychological								.15	.22	.55†	.55†
Eye									.38*	.67†	.66†
Mouth										.62†	.62†

* $P < .01$
† $P < .0001$

All test-retest correlations were at least .7 and ranged from .70 to .89. Correlations were ≥.80 for the energy, skin, psychological, and eye subscales and between .70 and .79 for the nutrition, lung, and mouth subscales. The test-retest correlation for the summary score was .79 for the full scale and .81 for the 28-item version. Compared with the original description of the instrument, test-retest scores were higher or the same for energy, skin, lung, psychological, eye, and summary scores and lower for nutrition and mouth but still adequate for all.

Interclass correlations showed that the energy subscale correlated with all subscales except the eye and mouth subscales. The mouth scores correlated with the nutrition and eye subscales, but overall the subscales were fairly independent. All except lung correlated with the summary score.

Table 4 shows that the LSS scores differed for each subscale ($P < .10$) and for the summary score ($P < .001$) between self-reported and NIH mild versus moderate/severe cGVHD except for 2 items. The lung subscale was not correlated with self-reported cGVHD severity, and the psychological scale was not correlated with the NIH cGVHD severity. Removal of the "need to use oxygen" item did not improve the results of the lung subscale analysis. In 2 other prospective multicenter observational studies conducted from 2007 to 2012 [14] and 2013 to 2017 [15], the rates of any endorsement of the "need to use oxygen" (3.2% and 6.1%) and "receiving nutrition from an IV or feeding tube" (1.2% and 2.2%) items were very low.

Table 4
cGVHD Symptoms by Self-Reported and NIH Calculated cGVHD Severity At Enrollment

	Self-Reported cGVHD Severity		
Symptoms	None*/Mild (n = 36)	Moderate/Severe† (n = 31)	P‡
Energy	19.0 (13.1)	38.1 (22.0)	<.001
Skin	9.2 (11.9)	21.6 (19.5)	.003
Nutrition	3.2 (8.0)	7.0 (8.6)	.070
Nutrition–modified§	4.1 (10.1)	8.7 (10.8)	.069
Lung	2.6 (4.4)	4.2 (7.7)	.315
Lung-modified§	3.0 (5.1)	4.9 (9.4)	.304
Psychological	12.3 (13.7)	21.4 (18.4)	.023
Eye	32.9 (26.3)	57.0 (25.8)	<.001
Mouth	14.6 (17.0)	25.8 (30.9)	.076
Summary	13.4 (7.9)	25.0 (9.6)	<.001
Summary-modified§	13.6 (8.0)	25.3 (9.9)	<.001
	cGVHD Severity per NIH Criteria		
	Mild (n = 10)	Moderate/Severe (n = 58)	P‡
Energy	11.1 (10.8)	30.9 (20.0)	.003
Skin	3.0 (3.5)	17.1 (17.5)	<.001
Nutrition	1.5 (3.4)	5.6 (8.9)	.014
Nutrition–modified§	1.9 (4.2)	7.0 (11.2)	.014
Lung	.0 (.0)	4.0 (6.6)	<.001
Lung–modified§	.0 (.0)	4.5 (7.8)	<.001
Psychological	10.0 (10.2)	17.7 (17.3)	.180
Eye	20.0 (16.8)	48.4 (28.2)	.003
Mouth	8.8 (13.2)	21.8 (26.1)	.024
Summary	7.8 (6.0)	20.8 (9.9)	<.001
Summary–modified§	7.8 (6.1)	21.0 (10.1)	<.001

Values are mean (standard deviation).

* Five patients indicated they had "none" and "0" or "1" severity on a 0-10 scale. They are included in the study because they had mild, moderate, or severe cGVHD per NIH criteria.

† One patient did not report cGVHD severity but was NIH severe so was grouped with the self-reported moderate/severe group.

‡ Based on *t*-test.

§ Modified from the original by deletion of 2 items (see text).

DISCUSSION

When the LSS was first developed and validated in the late 1990s, severe manifestations of cGVHD were more common. The survey also used a 1-month recall period because the intent was to use the instrument for clinical care and observational studies. This report shows that the items about need for oxygen or intravenous or tube feeding can be removed without adversely affecting test characteristics if a pure symptom scale is desired. The 7-day recall period may be used because the modified instrument retains its overall reliability and validity. Both changes result in a modified LSS (mLSS) that is better suited to clinical trials.

Results from this study and from reanalysis of 2 earlier cohorts show that endorsement of the oxygen and intravenous/feeding tube items was very infrequent. Although these questions were originally conceived to reflect bother due to the severity of cGVHD requiring need for such supportive care, they do not directly reflect cGVHD symptoms because even very symptomatic patients might refuse oxygen or feeding tubes. The low rate of endorsement seen in modern studies may also be due to better recognition and earlier/more effective treatment of cGVHD, although these hypotheses are speculative. Regardless, this study shows that these 2 items may be removed from the scoring algorithm. Although absolute scores will be higher because of removing items that are usually scored as zeros and bring down the average, as long as the enrollment and follow-up surveys are scored using the same formula, change scores are interpretable. Collection of the full 30-item version allows calculation of either the full or modified scale scores.

A previous study asked patients to compare how they would report their symptoms with a 7-day or 1-month time frame, showing that some patients reported the time frame selected would have altered their answers. The primary reason given was that their cGVHD symptoms had changed for better or worse in the past month, which is a legitimate reason for different answers, further justifying the change to a shorter recall period.

Intraclass correlations of 3 subscales (nutrition, lung, and psychological symptoms) were <.7, suggesting the items are not measuring a single construct. Examination of individual questions supports this conclusion; for example, the nutrition subscale includes difficulty swallowing, nausea, and weight loss, all recognized symptoms and signs of gastrointestinal cGVHD that are not always found together.

Limitations of this study include the modest sample size and restriction to outpatients from 1 center. Participants were well educated with 64% being college graduates. Patients with self-reported severe cGVHD were under-represented (3%), whereas there were 41% with severe cGVHD per the NIH criteria. Very few patients had liver and gastrointestinal symptoms, and patients were only stable or improved (none worsened) between the 2 test and retest measurements, which might be explained because the retest survey was administered only 1 week after a clinic visit where symptoms may have been detected and treated.

In summary, our results document the reliability and validity of the 7-day mLSS for evaluating cGVHD symptoms and suggest a 5- to 6-point difference in the summary score is clinically meaningful. The 7-day mLSS may be used in modern clinical trials.

ACKNOWLEDGMENTS

Financial disclosure: Supported by National Cancer Institute (grants CA118953 and CA163438).

Conflict of interest statement: There are no conflicts of interest to report.

REFERENCES

1. Lee SJ. Classification systems for chronic graft-versus-host disease. *Blood*. 2017;129:30–37.
2. 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.
3. Lee SJ, Wolff D, Kitko C, 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. The 2014 Response Criteria Working Group report. *Biol Blood Marrow Transplant*. 2015;21:984–999.
4. Pavletic SZ, Martin P, Lee SJ, 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.
5. Merkel EC, Mitchell SA, Lee SJ. Content validity of the Lee chronic graft-versus-host disease symptom scale as assessed by cognitive interviews. *Biol Blood Marrow Transplant*. 2016;22:752–758.
6. Lee SJ, Logan B, Westervelt P, et al. Comparison of patient-reported outcomes in 5-year survivors who received bone marrow vs peripheral blood unrelated donor transplantation: long-term follow-up of a randomized clinical trial. *JAMA Oncol*. 2016;2:1583–1589.
7. Walker I, Panzarella T, Couban S, et al. Pretreatment with anti-thymocyte globulin versus no anti-thymocyte globulin in patients with haematological malignancies undergoing haemopoietic cell transplantation from unrelated donors: a randomised, controlled, open-label, phase 3, multicentre trial. *Lancet Oncol*. 2016;17:164–173.
8. Miklos D, Cutler CS, Arora M, et al. Ibrutinib for chronic graft-versus-host disease after failure of prior therapy. *Blood*. 2017;130:2243–2250.
9. Williams KM, Cheng GS, Pusic I, et al. Fluticasone, azithromycin, and montelukast treatment for new-onset bronchiolitis obliterans syndrome after hematopoietic cell transplantation. *Biol Blood Marrow Transplant*. 2016;22:710–716.
10. Arai S, Pidala J, Pusic I, et al. A randomized phase II crossover study of imatinib or rituximab for cutaneous sclerosis after hematopoietic cell transplantation. *Clin Cancer Res*. 2016;22:319–327.
11. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ*. 2018;48:1273–1296.
12. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol*. 1994;47:81–87.
13. Lydick E, Epstein RS. Interpretation of quality of life changes. *Qual Life Res*. 1993;2:221–226.
14. Chronic GVHD Consortium. Rationale and design of the chronic GVHD cohort study: improving outcomes assessment in chronic GVHD. *Biol Blood Marrow Transplant*. 2011;17:1114–1120.
15. Chronic GVHD Consortium. Design and patient characteristics of the chronic graft-versus-host disease response measures validation study. *Biol Blood Marrow Transplant*. 2018;24:1727–1732.