Supplementary Appendix. Survey questions and answers

Diagnosis		
Question		*Answer*
Q1.	Currently, there are no diagnostic features of chronic GVHD for the eyes, liver and GI tract except esophagus. Should there be?	
	- No	58.7%
	- Yes	41.3%
Q2.	There are no distinctive features of chronic GVHD for the GI tract and liver. Should there be?	
	- No	63.0%
	- Yes	37.0%
Q3.	Should 2 or more distinctive manifestations be considered sufficient to diagnose chronic GVHD?	
	- Yes	85.1%
	- No	14.9%
Q4.	Should there be any pediatric modifications to the categorizations of diagnostic and distinct manifestations?	
	- No	69.8%
	- Yes	30.2%
Subcategory of GVHD		
Question		*Answer*
Q5.	Do patients with chronic GVHD have to have acute manifestations in all 3 sites (skin, liver and GI) to be called as “overlap” chronic GVHD?	
	- Yes. Patients need to have involvement in all of the 3 sites to be considered overlap.	4.3%
	- No. Only one site with acute manifestations is enough to diagnose overlap chronic GVHD.	95.7%
Q6.	Since the new classification is based on signs and symptoms, and not on time after transplant, should the terms “acute” and “chronic” be revised to something else?	
	- Yes	34.0%
	- No	66.0%
Q7.	How would you categorize the following patient? The patient had overlap chronic GVHD (oral lichenoid changes and gut GVHD). All manifestations were completely resolved after six months of systemic treatment. Now 2 years after transplantation, the patient has recurrent lower gut GVHD without any other signs of chronic GVHD.	
	- The patient has late acute GVHD (Type of GVHD is determined by the current condition, not by history of GVHD).	44.7%
	- The patient has overlap chronic GVHD (Once patients have chronic GVHD, they are never considered to have just acute GVHD again).	55.3%
Q8.	Do you think overlap chronic GVHD and progressive onset are interchangeable terms?	
	- Yes, they are the same.	17.8%
	- No. (please explain how they differ)	82.2%
Q9.	When you diagnose recurrent late acute GVHD or quiescent chronic GVHD, how many days of acute GVHD resolution are required before symptoms start again?	
	- At least 1 day	6.1%
	- At least 7 days	29.8%
	- At least 14 days	42.6%
	- Other	21.3%

Pathology		
Question		*Answer*
Q10.	In which sites can pathologists distinguish between acute and chronic GVHD? Check all that apply	
	- None	19.1%
	- Skin	74.5%
	- Mouth	61.7%
	- Esophagus	44.7%
	- GI tract except for esophagus	21.3%
	- Liver	12.8%
	- Lung	70.2%
	- Fascia	51.1%
	- Genital tract	44.7%
Q11.	Can pathologists confidently diagnose liver chronic GVHD?	
	- No	80.9%
	- Yes	19.1%
NIH organ score		
Question		*Answer*
Q12.	The current consensus recommends rating all symptoms even if you do not think the symptoms are due to GVHD (page 951). Should this be revised?	
	- Yes. Only symptoms that you attribute to GVHD should be scored.	21.3%
	- Yes. Symptoms that you unequivocally attribute to other etiologies should be excluded. If you are not sure of the etiology, they should be scored.	57.4%
	- Yes. I would downgrade the score by 1 if other etiologies contribute to symptoms (similar to acute GVHD grading).	8.5%
	- No. All symptoms should be scored regardless of attribution.	12.8%
NIH global score		
Question		*Answer*
Q13.	Performance status scoring is not incorporated into the NIH global scoring system. Should this be revised?	
	- Yes. It should be included in the global scoring system.	57.4%
	- No.	42.6%
Q14.1.	Should other manifestations (thrombocytopenia, eosinophilia, ascites, pleural effusion, etc. page 953) be included in the global scoring system?	
	- Yes.	80.9%
	- No.	19.1%
Q14.2.	If yes to above, which manifestation(s) should be included (page 953)? Check all that apply.	
	- Esophageal stricture or web	73.8%
	- Pericardial effusion	61.9%
	- Pleural effusion	61.9%
	- Ascites (serositis)	59.5%
	- Nephrotic syndrome	47.6%
	- Peripheral neuropathy	26.2%
	- Myathenia Gravis	42.9%
	- Cardiomyopathy	19.0%
	- Eosinophilia	59.5%
	- Polymyositis	59.5%
	- Cardiac conduction defects	16.7%
	- Coronary artery involvement	16.7%
	- Platelets <100,000/ul	66.7%

	- None of the above	7.1%
Response		
Question		*Answer*
Q15.	The current consensus recommends rating organ severity without distinguishing between active disease and fixed deficits such as skin sclerosis, dry eyes, bronchiolitis obliterans, joint contracture and vaginal stenosis (page 951). Should this be revised? - Yes, it should be revised for both severity scoring and response measurement - Yes, it should be revised only for response measurement. - No.	 43.8% 39.6% 16.7%
Skin		
Question		*Answer*
Q16.	Skin biopsy is not mandatory for diagnosis of skin chronic GVHD. Should this be revised? - Yes. Skin biopsy should be mandatory. - No.	 18.8% 81.3%
Q17.	If you see hidebound changes in only a small area (for example 1% of legs), the NIH skin score is 3. Should this be revised? - Yes. - No.	 70.8% 29.2%
Q18.	The current consensus recommends rating organ severity without distinguishing between active disease and fixed deficits. Thus if skin manifestations are not active (for example, hyperpigmentation after resolution of an erythematous rash), they are currently included in calculating affected body surface area (BSA). Should this be revised? - Yes. - No.	 79.2% 20.8%
Q19.	Various skin manifestations are currently listed in the scoring sheet but which of them should be considered when calculating BSA? Check all that apply. - Maclopapular rash - Lichen planus-like feature - Papulosquamous lesions or ichthyosis - Hyperpigmentation - Hypopigmentation - Keratosis pilaris - Erythema - Erythroderma - Poikiloderma - Sclerotic features - Pruritus - Hair involvement - Nail involvement	 95.8% 95.8% 85.4% 47.9% 50.0% 68.8% 91.7% 95.8% 81.3% 95.8% 16.7% 60.4% 52.1%
Q20.	Is just nail and/or hair involvement sufficient for NIH skin score 1? - Yes. - No.	 72.9% 27.1%
Q21.	Is just pruritus (without any skin changes) sufficient for NIH skin score 1 or greater? - Yes. - Usually not, but I would score it if pruritus is severe. - No, in any case.	 12.5% 41.7% 45.8%
Q22.	Is just hyperpigmentation and/or hypopigmentation of skin sufficient for NIH skin score 1 or greater? - Yes. I always include them in calculating body surface area (BSA).	 41.7%

	- Yes. I include them only in the baseline assessment. - No. I never include them in calculating BSA.	22.9% 35.4%
Q23.	How can you distinguish deep from superficial sclerosis in the abdominal skin among obese patients? - Difficult to distinguish - Possible to estimate (by what means?)	 83.3% 16.7%
Q24.	Are there any imaging methods that can distinguish deep from superficial sclerosis? - No - Yes	 67.4% 32.6%
Q25.	When you see hyperpigmentation and lichenoid in the same area, how do you grade and assign involved body surface area in the Total Skin Score (Vienna Score)? score 1=discolored; score 2=lichenoid plaque or skin thickened (able to move) - Grade total area as score 2. - Grade half area as score 1 and half as score 2.	 91.7% 8.3%
Eye		
Question		*Answer*
Q26.	A patient had moderate ocular dryness very early after transplantation (for example, 30 days after transplantation), with no other manifestations of chronic GVHD. Should it be diagnosed as acute GVHD, chronic ocular GVHD or neither? - Acute GVHD. - Chronic GVHD. - Neither acute nor chronic GVHD.	 12.5% 8.3% 79.2%
Q27.	A patient has just been diagnosed with chronic GVHD in the mouth. If the patient has been using eye drops 3 times a day due to dry eye starting before transplant, should the patient be scored for ocular chronic GVHD? - Yes. - No.	 12.5% 87.5%
Q28.	If a patient had punctual plugging and had symptomatic relief to such an extent that he requires eye drops only 2 times a day, how should we determine the NIH eye score when you are sure that punctual plugs are still in the eyes? - Score 1 (Mild dry eye symptoms not affecting ADL (requiring eyedrops ≤3 x per day) OR asymptomatic signs of keratoconjunctivitis sicca) - Score 2 (Moderate dry eye symptoms partially affecting ADL (requiring drops > 3 x per day or punctal plugs), WITHOUT vision impairment)	 27.1% 72.9%
Q29.	The same scenarios as above. How would you rate the NIH eye score after the punctal plugs fall out if he still uses eyedrops 2 times a day? - Score 1 - Score 2	 87.0% 13.0%
Q30.	If a patient started special contact lenses for treatment of ocular GVHD and had symptomatic relief to such an extent that he requires eye drops 2 times a day, how should we rate the NIH eye score - Score 1 - Score 2 - Score 3	 17.0% 21.3% 61.7%
Q31.	A patient uses eye drops 2 times a day due to dry eye prior to transplant. After transplant, he is diagnosed with chronic GVHD and increases the frequency of eye drops to 4 times a day due to worsening eye dryness. How do you rate the NIH eye score? - Score 0 - Score 1 - Score 2	 2.2% 50.0% 47.8%
Q32.	Should all types of eye drops be included when counting the frequency of eye drops for rating the NIH eye score (i.e., also include cyclosporine eye drop, steroid eye	

	drop, antibiotic eye drops)? - Yes. - No.	 34.8% 65.2%
Q33.	If a patient has “new ocular sicca documented by Schirmer test” or “a new onset of keratoconjunctivitis with Schirmer score <10 mm”, is this sufficient to diagnose chronic GVHD? - Yes, even if only eyes are involved. - No, a distinctive manifestation in 1 other site is required to diagnose chronic GVHD.	 41.7% 58.3%
Q34.	Should we consider excessive tearing as one form of GVHD in the NIH eye score? - Yes. - No.	 56.3% 43.8%
Q35.	If eye symptoms differ between left and right eyes, which eye do you use for rating the NIH eye score? - The worse eye - The better eye - The average of the two eyes	 83.3% 0% 16.7%
Q36.	If a patient lost vision in one eye because of chronic GVHD but is completely asymptomatic in the other eye, what is the NIH eye score? - Score 0 - Score 3 - Other	 2.1% 87.5% 10.4%
Mouth		
Question		*Answer*
Q37.	If a patient has diagnostic signs such as lichenoid changes but has no oral symptoms, the NIH mouth score is 0. Should this be revised? - Yes. It should be score 1. - No.	 77.1% 22.9%
Q38.	How should we rate the NIH mouth score for moderate oral sensitivities without lichenoid changes or other signs of chronic GVHD? - Score 0 - Score 1 (mild symptoms with disease signs but not limiting oral intake significantly) - Score 2 (moderate symptoms with disease signs with partial limitation of oral intake)	 22.9% 54.2% 22.9%
Q39.	If the patient has extensive oral lichenoid changes but has only mild symptoms, how should we rate the NIH mouth score? - Score 1 - Score 2 - Score 3 (Severe symptoms with disease signs on examination with major limitation of oral intake)	 47.9% 50.0% 2.1%
Q40.	Should gingivitis, oral mucositis and pain continue to be considered “common” signs, even if mouth is not a recognized target organ in acute GVHD? Or should these be considered distinctive signs of chronic GVHD? - They should be remained as common signs. - They should be considered distinctive signs for chronic GVHD.	 40.4% 59.6%
Q41.	Should we consider superficial mucoceles that come and go when you determine the NIH mouth score? - Yes, if they are present. - Yes, if they are symptomatic. - No, mucoceles should not be considered at all. - Other	 43.5% 21.7% 32.6% 2.2%

Gastrointestinal tract		
Question		*Answer*
Q42.	Should we score diarrhea in the NIH GI scoring section when the patient has chronic GVHD in other sites but biopsy is negative for GI GVHD? - Yes. - No.	 66.0% 34.0%
Q43.	How do you score diarrhea if no biopsy is done but a patient has diagnostic chronic GVHD in other sites? - Score diarrhea as GI GVHD. - Score 0 unless a biopsy shows GI GVHD. - Other	 60.4% 14.6% 25.0%
Q44.	A patient has chronic GVHD in other sites plus nausea and anorexia. Is a biopsy required for the diagnosis of GI involvement? - Yes, a biopsy is always needed to diagnose GI involvement. - No, nausea and anorexia without biopsy are sufficient to diagnosis GI GVHD if a patient has chronic GVHD in other organs.	 45.8% 54.2%
Q45.	What is the reference time point used for calculation of weight loss? - 1 month ago - 3 months ago - 6 months ago - 12 months ago - Pre-transplant - Other	 44.7% 20.2% 8.5% 2.1% 12.8% 11.7%
Q46.	A patient has mild loose stool and the colon biopsy is positive for GVHD. The patient also had 10% weight loss as compared to one month ago, but you attribute the weight loss to poorly controlled steroid-induced diabetes. How do you rate the NIH GI score for this patient? - Score 1 (symptoms such as dysphagia, anorexia, nausea, vomiting, abdominal pain or diarrhea without significant weight loss (<5%)) - Score 2 (symptoms associated with mild to moderate weight loss (5-15%))	 36.2% 63.8%
Liver		
Question		*Answer*
Q47.	What laboratory tests should be considered for staging of late acute GVHD in the liver? - Only bilirubin is considered (the same staging criteria for classic acute GVHD are applied to late acute GVHD). - ALT, AST, bilirubin and alkaline phosphatase are considered (the same staging criteria for NIH chronic GVHD are applied to late acute GVHD).	 31.9% 68.1%
Q48.	A patient has normal AST, ALT and bilirubin, but has elevated alkaline phosphatase (AP).You do not have isozyme information for AP. How do you rate the NIH liver score? - Score 0 - Score according to the level of AP elevation. - Check AP isozymes. If it is bone origin, score 0. If it is liver origin, score according to the level of elevation. - Check gamma GTP. If it is elevated, attribute AP elevation to liver and score accordingly. - Other	 25.0% 20.8% 29.2% 18.8% 6.3%
Q49.1.	If a patient has already been diagnosed with chronic GVHD of the mouth and	

	eyes, and also has LFT abnormalities but no liver biopsy, should the LFT abnormalities be considered GVHD? - Yes. - No. You always need liver biopsy to confirm GVHD.	 83.0% 17.0%
Q49.2	If yes to above, how do you classify the patient? - Overlap chronic GVHD. - Classic chronic GVHD, since liver biopsy is not confirmed.	 46.5% 53.5%
Q50.	If other etiologies are confirmed for liver abnormalities (for example, hemochromatosis, viral hepatitis, leukemia invasion, drug side effect, or alcohol consumption), how do you rate the NIH liver score for patients with chronic GVHD? - Score 0 - Score according to LFT abnormalities - Decrease the score by 1 point as with acute GVHD staging. - Other	 59.4% 20.8% 15.6% 4.2%
Joint/Fascia/Muscle		
Question		*Answer*
Q51.	How should we rate joint tightness due to prior injury or avascular necrosis in the NIH joint score? - Score 0. - Score according to functional deficit without regard to etiology.	 82.6% 17.4%
Q52.	Is isolated early fasciitis manifested by edema diagnostic for chronic GVHD? - Yes, always. - Yes, but requires confirmation by imaging (MRI etc.) or biopsy. - No.	 32.3% 41.7% 26.0%
Q53.	How do you distinguish joint problems related to chronic GVHD from other cause of joint impairment? - Careful history taking and physical examination of fascia and joints. - It is impossible to distinguish. Score regardless of etiology.	 87.5% 12.5%
Q54.	Do you think joint pain mimicking rheumatoid arthritis after transplant is joint GVHD? - Yes. - No. - I don't know.	 52.1% 8.3% 39.6%
Q55.	If muscle biopsy is positive for myositis but diagnostic or distinctive features of chronic GVHD are absent in other sites, should it be sufficient to diagnose chronic GVHD? - Yes. - No.	 79.2% 20.8%
Lung		
Question		*Answer*
Q56.	If a patient has no diagnostic or distinctive chronic GVHD in other sites, chronic GVHD cannot be diagnosed even if a patient has obstructive changes on pulmonary function test (PFT) and air trapping, small airway thickening and bronchiectasis on CT (i.e., clinical bronchiolitis syndrome). Should this be revised? - Yes, clinical bronchiolitis obliterans syndrome should be considered a diagnostic manifestation (sufficient to make the diagnosis of chronic GVHD). - No, clinical bronchiolitis obliterans syndrome should remain a distinctive manifestation, requiring confirmation in another organ before the diagnosis of chronic GVHD applies.	 70.8% 29.2%
Q57.	Is cryptogenic organizing pneumonia a form of lung GVHD?	

	-Yes. - No.	57.4% 42.6%
Q58.	Should cryptogenic organizing pneumonia still be considered a “common” sign, since lung is not a recognized target organ in acute GVHD? - Yes, COP should remain a common sign. - No, COP should be considered a distinctive form of chronic GVHD. - No, COP should be removed as a manifestation of GVHD.	 38.3% 31.9% 29.8%
Q59.	How do you score dyspnea that you believe is due to steroid myopathy? - Score 0. - Score according to symptoms.	 72.3% 27.7%
Q60.	When lung symptom scores differ from PFT scores, higher values are used for final lung scores. Should we always use PFT scores to determine the lung score whenever PFT results are available even if the patient has no pulmonary symptoms? - Yes. PFT scores should remain a part of the scoring system regardless of cause or symptoms. - No. PFT scores should only be used if a patient has pulmonary symptoms and/or bronchiolitis obliterans.	 66.7% 33.3%
Q61.	If a patient has pulmonary symptoms with normal PFT, how should we score the NIH lung score? - Score 0, since you are not sure whether the symptoms are due to GVHD. - Score according to symptoms. - I would downgrade the score by 1.	 66.7% 27.1% 6.3%
Q62.	If a patient has chronic obstructive pulmonary disease (COPD) before transplant, how should we determine the NIH PFT score after transplant? - Score 0, since you attribute abnormal PFTs to COPD. - Score according to post transplant PFTs, without consideration of pretransplant PFTs. - Score according to post transplant PFTs only if they worsen from pretransplant PFTs. - Score according to post transplant PFTs, but downgrade by 1.	 4.2% 4.2% 89.6% 2.1%
Q63.	If a patient doesn't have PFTs results, how should we determine the NIH lung score? - Score according to symptoms. - Leave NIH lung score blank (missing). - Other	 56.3% 37.5% 6.3%
Genital tract		
Question		*Answer*
Q64.	Is erectile dysfunction a symptom of genital GVHD? - Yes - No	 8.5% 91.5%
Q65.	Should we score the genitals for men? - Yes - No	 63.8% 36.2%
Q66.	Can the NIH genital score be completed without gynecological exam? - Yes. If there is no gynecologic exam, score according to symptoms. - No, a gynecological exam is required. - Other	 43.8% 43.8% 12.5%
Q67.	A female patient is asymptomatic due to sexual inactivity and has moderate signs of genital GVHD. What is the NIH genital score? - Score 0 (asymptomatic) - Score 1 (symptomatic with mild signs on exam AND no effect on coitus and minimal discomfort with gynecologic exam)	 8.5% 40.4%

	- Score 2 (symptomatic with moderate signs on exam AND with mild dyspareunia or discomfort with gynecologic exam)	51.5%
Q68.	A female patient tells you that she is asymptomatic but uses a dilator for her fixed moderate vaginal stricture. What is the NIH genital score? - Score 0 - Score 1 - Score 2 - Score 3 (symptomatic WITH advanced signs (stricture, labial agglutination or severe ulceration) AND severe pain with coitus or inability to insert vaginal speculum)	 2.1% 19.1% 31.9% 46.8%
Q69.	Is vaginal dryness sufficient for NIH genital score 1 or greater? - Yes, for all women. - Yes, but only for pre-menopausal women. - No.	 16.7% 35.4% 47.9%
Q70.	What is the NIH genital score for a female who has mild dyspareunia and severe signs on gynecological exam? - Score 2 - Score 3	 31.3% 68.8%
Other sites		
Question		*Answer*
Q71.	How do you determine if peripheral neuropathy is due to chronic GVHD in a patient with an established diagnosis of chronic GVHD? - Attribute neuropathy to chronic GVHD if it is new or worsening without other cause - You can never tell.	 45.8% 54.2%
Q72.	Should nephrotic syndrome after allogeneic transplantation be considered a sign of chronic GVHD? - Yes, always. - Yes, only in a patient with an established diagnosis of chronic GVHD. An Isolated nephrotic syndrome is not chronic GVHD. - No, in any case.	 15.2% 65.2% 19.6%