Biology of Blood and Marrow Transplantation

journal homepage: www.bbmt.org

ASBMT
American Society for Blood
and Marrow Transplantation

Report

International, Multicenter Standardization of Acute Graft-versus-Host Disease Clinical Data Collection: A Report from the Mount Sinai Acute GVHD International Consortium

CrossMark

Andrew C. Harris [1,2], Rachel Young [1,3], Steven Devine [4], William J. Hogan [5], Francis Ayuk [6], Udomsak Bunworasate [7], Chantiya Chanswangphuwana [7], Yvonne A. Efebera [4], Ernst Holler [8], Mark Litzow [5], Rainer Ordemann [9], Muna Qayed [10], Anne S. Renteria [3], Ran Reshef [11], Matthias Wölfl [12], Yi-Bin Chen [13], Steven Goldstein [1], Madan Jagasia [14], Franco Locatelli [15], Stephan Mielke [16], David Porter [17], Tal Schechter [18], Zhanna Shekhovtsova [19], James L.M. Ferrara [3], John E. Levine [1,3,*]

[1] *Blood and Marrow Transplantation Program, University of Michigan, Ann Arbor, Michigan*
[2] *Blood and Marrow Transplantation Program, University of Utah, Salt Lake City, Utah*
[3] *Blood and Marrow Transplantation Program, The Icahn School of Medicine at Mount Sinai Hospital, New York, New York*
[4] *Blood and Marrow Transplantation Program, Ohio State University, Columbus, Ohio*
[5] *Blood and Marrow Transplantation Program, Mayo Clinic, Rochester, Minnesota*
[6] *Department of Stem Cell Transplantation, University Medical Center, Hamburg-Eppendorf, Germany*
[7] *Blood and Marrow Transplantation Program, Chulalongkorn University, Bangkok, Thailand*
[8] *Blood and Marrow Transplantation Program, University of Regensburg, Regensburg, Germany*
[9] *Blood and Marrow Transplantation Program, University Hospital TU Dresden, Dresden, Germany*
[10] *Pediatric Blood and Marrow Transplantation Program, Aflac Cancer and Blood Disorders Center, Emory University and Children's Healthcare of Atlanta, Atlanta, Georgia*
[11] *Blood and Marrow Transplantation Program, Columbia University Medical Center, New York, New York*
[12] *Pediatric Blood and Marrow Transplantation Program, Children's Hospital, University of Würzburg, Würzburg, Germany*
[13] *Bone Marrow Transplantation Program, Massachusetts General Hospital, Boston, Massachusetts*
[14] *Division of Hematology-Oncology, Vanderbilt University Medical Center, Nashville, Tennessee*
[15] *Pediatric Blood and Marrow Transplantation Program, Ospedale Pediatrico Bambino Gesu, Rome, Italy*
[16] *Blood and Marrow Transplantation Program, University of Würzburg, Würzburg, Germany*
[17] *Blood and Marrow Transplantation Program, University of Pennsylvania, Philadelphia, Pennsylvania*
[18] *Pediatric Blood and Marrow Transplantation Program, The Hospital for Sick Children, Toronto, Ontario, Canada*
[19] *Federal Clinical Research Center for Children's Hematology, Oncology and Immunology, Moscow, Russian Federation*

ABSTRACT

Acute graft-versus-host disease (GVHD) remains a leading cause of morbidity and nonrelapse mortality after allogeneic hematopoietic cell transplantation. The clinical staging of GVHD varies greatly between transplant centers and is frequently not agreed on by independent reviewers. The lack of standardized approaches to handle common sources of discrepancy in GVHD grading likely contributes to why promising GVHD treatments reported from single centers have failed to show benefit in randomized multicenter clinical trials. We developed guidelines through international expert consensus opinion to standardize the diagnosis and clinical staging of GVHD for use in a large international GVHD research consortium. During the first year of use, the guidance followed discussion of complex clinical phenotypes by experienced transplant physicians and data managers. These guidelines increase the uniformity of GVHD symptom capture, which may improve the reproducibility of GVHD clinical trials after further prospective validation.

© 2016 American Society for Blood and Marrow Transplantation.

INTRODUCTION

The practice of acute graft-versus-host disease (GVHD) clinical staging varies between transplant centers, and poor concordance has been recognized for over 25 years, with

agreement between independent reviewers ranging from 40% to 72% [1,2]. The threshold for diagnosis of GVHD and quantification of symptoms remain variable among centers, as evidenced by widely disparate cumulative incidences of grades II to IV GVHD ranging from 40% to 80% after T cell–replete blood and marrow transplantation (BMT) [3,4]. The difference in rates are likely due to multiple factors such as frequency of obtaining target organ biopsies when symptoms arise, appropriate consideration and application of measures to exclude alternative diagnoses, and absence of consensus guidelines to address inherent challenges in GVHD staging such as stool volume quantification. These variations in practice also result in differences in reporting of timing of GVHD onset and its severity, which poses a significant barrier to the successful conduct of multicenter trials. These barriers can only be overcome through the consistent application of guidelines to clinical GVHD data capture that can uniformly be applied by both clinicians and data managers.

We recently reported a new web-based remote data entry system we developed for GVHD data capture that is now being used in an international GVHD research consortium (the Mount Sinai Acute GVHD International Consortium, or MAGIC) [5]. The GVHD staging guidance outlined in this report was developed initially at a single institution (University of Michigan) and then tested in this multicenter group for clarity and ease of use. The guidance was then refined during its first year of use through discussion of complicated cases seen at these centers via international webinars to reach a consensus opinion. The guidelines outlined in this report are not intended to dictate clinical practice, but rather they are a tool to standardize the collection of complex clinical data for acute GVHD clinical and translational research. Although a few controversial topics remain to be resolved in this evolving guidance, the centers in our GVHD consortium have found this guidance for GVHD data collection clearer and easier to use than prior systems, and it could be used by other centers that conduct GVHD clinical research.

TARGET ORGAN SYMPTOM CAPTURE

Although the guidelines for GVHD staging by symptom severity have been well established [6], quantification of the severity of symptoms has not been standardized across centers. Many databases only record target organ staging, which limits retrospective review of severity without examining source documents from the medical record. To avoid this problem, we collect the absolute quantification of symptoms (extent of skin rash, total bilirubin level, volume of diarrhea), according to the guidance detailed below, regardless of suspected/proven etiology.

Skin

The skin is the most commonly involved GVHD target organ [7], but patients have multiple potential causes for rash after allogeneic transplantation (eg, GVHD, medications, viral exanthemata), and the clinical manifestations may not point to 1 specific etiology. Rash quantification can be clinically challenging, however, because inclusion of all areas of abnormal skin does not distinguish areas of active inflammatory erythema that are characteristic of GVHD [8] from areas of inactive hyperpigmentation or other non-GVHD changes that may lead to discrepancies in rash quantification from 1 clinician to the next. This distinction is important because the natural course of an erythematous rash while

resolving is to appear hyperpigmented or "browned over." For example, a patient may have rash involving 60% body surface area (BSA) that if included in its entirety would be classified as stage 3 skin GVHD. Upon closer inspection, however, the skin changes may include petechiae, hyperpigmentation, and/or other changes not consistent with active GVHD. If the non-GVHD skin changes account for more than 10% of the skin changes in this example, the GVHD would be downgraded to stage 2 skin GVHD, which may impact the decision to initiate systemic GVHD therapy. Thus, only areas involved with active erythema should be used for determination of BSA staging of GVHD using the "rule of nines" [9].

A portion of a body area segment may be used for the quantification. For example, if erythema is observed only on the upper half of an arm, this would be quantified as 4.5%, or half of the arm's total 9% BSA. Additionally, one should report if desquamation or fluid-filled bullae are present, because these findings are the hallmark of stage 4 skin GVHD. Patients may have occasional blisters or small patches of dry skin with desquamation that do not reflect the massive inflammation implied by stage 4 skin GVHD. Therefore, this guidance requires both generalized erythema as well as >5% BSA involvement with blisters and/or desquamation to diagnose skin stage 4 GVHD.

Liver

The liver is the least frequently involved acute GVHD target organ [4,10,11]; however, it is important to document the presence of liver GVHD because of the poorer prognosis it portends [12]. Liver GVHD staging is based solely on total (not conjugated/direct) serum bilirubin levels. Liver GVHD manifesting as transaminitis without hyperbilirubinemia is not staged when applying modified Glucksberg GVHD staging [6,13,14], and transaminitis from non-GVHD causes is common after BMT. We therefore only diagnose liver GVHD manifesting as transaminitis without concomitant elevation in serum bilirubin when the presence of GVHD is confirmed by liver biopsy and score it as stage 0. Future revisions to staging acute liver GVHD presenting as isolated transaminitis will require correlation of transaminase levels, the diagnoses under consideration, and treatment decisions with clinical outcomes.

Because of the relatively infrequent involvement of the liver at GVHD onset and the fact that patients may have hyperbilirubinemia from other causes at the onset of GVHD (eg, chemotherapy toxicity, sinusoidal obstructive syndrome, parenteral nutrition–associated cholestasis), if bilirubin levels were elevated before the diagnosis of GVHD in another target organ and do not increase further, we do not diagnose liver GVHD in the absence of biopsy confirmation. If hyperbilirubinemia develops at the same time or after the onset of GVHD in another target organ, however, liver GVHD is presumed to be present in the absence of an identified alternative cause.

Upper Gastrointestinal Tract

The frequency of GVHD involving the upper gastrointestinal (GI) tract is highly variable between BMT centers, with incidences ranging from 24% to 60% [3,15]. Symptoms of concern for upper GI GVHD include anorexia, nausea, vomiting, and dyspepsia [16], but these can also be seen frequently as a result of infection, mucositis, conditioning regimen toxicity, and/or medication side effect. The timing, severity, and duration of symptoms sufficient for clinical

Table 1
GVHD Target Organ Staging

Stage	Skin (Active Erythema Only)	Liver (Bilirubin)	Upper GI	Lower GI (stool output/day)
0	No active (erythematous) GVHD rash	<2 mg/dL	No or intermittent nausea, vomiting, or anorexia	Adult: <500 mL/day or <3 episodes/day Child: <10 mL/kg/day or <4 episodes/day
1	Maculopapular rash <25% BSA	2-3 mg/dL	Persistent nausea, vomiting or anorexia	Adult: 500-999 mL/day or 3-4 episodes/day Child: 10-19.9 mL/kg/day or 4-6 episodes/day
2	Maculopapular rash 25-50% BSA	3.1-6 mg/dL		Adult: 1000-1500 mL/day or 5-7 episodes/day Child: 20-30 mL/kg/day or 7-10 episodes/day
3	Maculopapular rash >50% BSA	6.1-15 mg/dL		Adult: >1500 mL/day or >7 episodes/day Child: >30 mL/kg/day or >10 episodes/day
4	Generalized erythroderma (>50% BSA) *plus* bullous formation and desquamation >5% BSA	>15 mg/dL		Severe abdominal pain with or without ileus or grossly bloody stool (regardless of stool volume).

Overall clinical grade (based on most severe target organ involvement):
Grade 0: No stage 1-4 of any organ.
Grade I: Stage 1-2 skin without liver, upper GI, or lower GI involvement.
Grade II: Stage 3 rash and/or stage 1 liver and/or stage 1 upper GI and/or stage 1 lower GI.
Grade III: Stage 2-3 liver and/or stage 2-3 lower GI, with stage 0-3 skin and/or stage 0-1 upper GI.
Grade IV: Stage 4 skin, liver, or lower GI involvement, with stage 0-1 upper GI.

concern for GVHD vary between centers, and assessment depends on close attention to caloric intake and symptom reporting. Although upper GI biopsy is required by modified Glucksberg staging [6], in practice this requirement is often ignored and could result in under-reporting of the incidence of upper GI GVHD. Although we encourage clinicians to obtain upper GI endoscopy whenever possible, we allow for the diagnosis of upper GI GVHD without biopsy confirmation provided the symptoms meet consensus thresholds for severity and duration to avoid overdiagnosis of upper GI GVHD that could occur without a biopsy requirement; we believe this is of particular importance when patients present without GVHD symptoms in additional target organs. We do not consider GVHD as a possible etiology when nausea lasts fewer than 3 days, or with fewer than 2 vomiting episodes per day for at least 2 days, or anorexia without weight loss.

After making a diagnosis of upper GI GVHD, the clinician must inquire and document whether symptoms persist, because these symptoms are very subjective in nature and must no longer be present to state upper GI GVHD has resolved. If a patient has lost weight as a result of upper GI GVHD, his or her weight must be stable or increasing—and not attributable to fluid overload—before upper GI can be considered quiescent.

Lower GI Tract

Lower GI GVHD is the target organ involvement most associated with nonrelapse mortality [3,4,7,10,17,18]. Staging of this target organ relies on accurate measurement of daily stool volumes and documentation of the presence of hematochezia or severe abdominal pain [6]. Post-BMT diarrhea often develops in the outpatient setting where stool volumes cannot be accurately recorded, and stool output is not always closely measured when patients are in the hospital, either because patients did not save the stool for quantification or because stool and urine were not collected separately. To address these barriers, we reviewed hospital flowsheets from 300 patients with post-transplant diarrhea from both GVHD and non-GVHD causes that contained both the measured volume and number of episodes of diarrhea. From these source documents we calculated an average volume of 200 mL per diarrhea episode for adults or 3 mL/kg for

children <50 kg (Table 1). Formed or mostly formed stools should not be quantified or counted toward diarrhea volume.

To stage lower GI GVHD at onset, we use the highest daily diarrhea volume during the 3 days before its diagnosis, after excluding diarrhea volumes attributable to bowel preps for endoscopy procedures. Because stool volumes may vary widely from day to day and in response to treatment, after initiation of treatment we stage lower GVHD based on the diarrhea volume using (in the order of preference) (1) average of 3 consecutive days, (2) average of 2 consecutive days, or (3) the volume on day of assessment.

Finally, severe abdominal pain, ileus, and/or grossly bloody stool should be documented when present because stage 4 lower GI GVHD is staged based on the presence of these symptoms and is independent of volume of diarrhea [19]. We do not consider streaks of blood in the stool due to hemorrhoids or anal fissures or transient hematochezia after endoscopic biopsies when making this determination. The definition of severe abdominal pain remains a judgment call on the part of the treating clinician, but we suggest only considering pain attributed to GVHD that requires the initiation of high doses of narcotic pain medication or a significant increase in ongoing narcotic use *and* the abdominal pain significantly impacts a patient's performance status as determined by the treating clinician.

CONFIDENCE LEVELS

Thresholds for initiation of systemic corticosteroids for treatment of GVHD are variable among centers, particularly in the setting when no biopsy has been obtained and/or when only limited skin GVHD is present. These differences in GVHD diagnosis make it challenging to conduct multicenter trials because of the difficulty determining whether a patient truly experienced GVHD. To address these varying practices, we developed a structure for collecting granular GVHD data and assigning confidence levels for the attribution of symptoms to GVHD based on the treatment decisions made by the clinician. When symptoms develop concerning for GVHD, a confidence level is assigned to the diagnosis of GVHD as supported by categorized facts (eg, biopsy results and infectious studies) and by clinical action (Table 2). These confidence levels permit reproducible data reporting and may facilitate selection of patients for research studies and do not

Table 2
Confidence Level Criteria

Confidence Level	Pathologic Evidence	Clinician Assessment	Treatment for Acute GVHD	Comments
Confirmed	Unequivocal pathologic evidence of GVHD	GVHD is the etiology for symptoms	Not applicable	GVHD is clearly present even if other etiologies may coexist simultaneously.
Probable	Not required	GVHD most likely etiology for symptoms (as evidenced by treatment being provided)	Yes	GVHD is most likely present, but other etiologies may also explain the symptoms, and there is insufficient evidence to make a confirmed diagnosis.
Possible	Not required	GVHD in differential diagnosis (but no treatment is being provided)	No	GVHD may be present, but other etiologies are favored to the degree that GVHD treatment is not initiated.
Negative	Unequivocal evidence of a diagnosis other than GVHD (eg, drug rash)	GVHD is not considered as an explanation for the symptoms	No and the symptoms resolve without GVHD treatment	A "negative" biopsy (eg, normal skin) is not unequivocal evidence of a diagnosis other than GVHD.

imply validated criteria of confidence. A *negative* confidence level is assigned when either there are no active GVHD symptoms or an alternative etiology is present and treated (if indicated) and GVHD is not entertained as a possible diagnosis. GVHD is deemed *possible* when a diagnosis of GVHD is under consideration but not confirmed by biopsy and the clinical suspicion for GVHD is not sufficiently high that the treating physician initiated GVHD therapy. A confidence level of *probable* reflects that treatment for GVHD was started because of sufficient clinical concern but the diagnosis is not confirmed by biopsy, either because no biopsy was obtained or because GVHD is not clearly identified on a biopsy. Finally, GVHD is considered *confirmed* if there is unequivocal evidence of GVHD on a biopsy.

Standardization of the determination of GVHD onset is extremely important for clinical research because onset is typically reported as a "time-to-event" variable. Using these confidence levels, GVHD onset is determined to be the date of initiation of therapy for GVHD or confirmation of GVHD by biopsy, whichever is earlier. We do not back-date the onset of GVHD to the onset of symptoms subsequently diagnosed as GVHD because we cannot exclude the possibility that symptoms were attributable to an alternative cause when they first appeared only to have GVHD subsequently develop. The scenario of isolated stage 1 to 2 skin GVHD represents a special case because many clinicians do not obtain biopsies or start systemic GVHD therapy, and clinicians will often start topical steroid treatment for limited rashes of any cause, suggesting a lower clinical threshold for initiating topical therapy than for starting systemic corticosteroid therapy. The specificity of skin biopsies for GVHD is also notoriously low, and pathologic findings often overlap between GVHD, drug reaction, and other causes. To standardize onset determination in this scenario, we consider the date on which a clinician first clinically diagnoses a rash as GVHD as onset. Because GVHD rarely occurs before day 14 post-transplant [10] and symptoms such as rash, nausea, and diarrhea are often present, a confidence level of "possible" is not entertained in this early time period. This approach decreases what would otherwise be many false GVHD diagnoses because many clinicians will consider even highly unlikely events "possible" if not required to act on them. However, when GVHD is diagnosed and treated, a confidence level of probable or confirmed can be assigned during the first 2 weeks post-transplant depending on whether GVHD is biopsy-proven.

BIOPSY INTERPRETATION

Biopsies are often obtained to confirm a GVHD diagnosis, but experienced pathologists from different centers disagree on the threshold of histopathologic findings that should be present to diagnose acute GVHD [20]. Biopsy interpretation can be further complicated by the timing of the biopsy post-transplant and by the setting in which the symptoms arise and may not clearly identify the etiology of GVHD-like symptoms in up to 60% of biopsies [21,22]. These inconsistencies can result in highly variable treatment decision-making among clinicians, and just as we do not attempt to influence clinical decision-making with these research data collection guidelines, we do not intend to dictate the clinical practice of our colleagues in pathology.

Because biopsy reports may be highly variable in their language, we have classified the interpretation of biopsy reports into 4 categories: *non-GVHD etiology* (unequivocal evidence of a diagnosis other than GVHD without concomitant features suggestive of GVHD), *nondiagnostic* (no abnormalities identified, or subtle changes noted that are insufficient to identify an etiology, or insufficient tissue for interpretation), *equivocal* (findings are present consistent with GVHD and other etiologies but the pathologist cannot definitively confirm the presence of GVHD), and *positive* (GVHD is clearly identified as present on the biopsy, with or without other coexisting processes present on the biopsy). Biopsy results are used together with the clinical treatment decisions to assign a confidence level to the attribution of symptoms to a diagnosis of GVHD (Table 3).

Table 3
Biopsy Results and Confidence Levels

Pathology Results	Target Organ Confidence Level		
	Treated as GVHD	Not Treated but GVHD in Differential Diagnosis	Not Treated and GVHD Not in Differential Diagnosis
Positive	Confirmed	Confirmed	Confirmed
Equivocal	Probable	Possible	Possible
Nondiagnostic	Probable	Possible	Negative
Non-GVHD etiology	Probable	Negative	Negative

GVHD confirmed in a biopsied target organ raises the confidence level from possible to probable for other target organs where GVHD is suspected, even in the absence of treatment.

At times, a biopsy report lacking clear language remains difficult to definitively categorize, which necessitates discussion and adjudication on a multicenter call to obtain consensus opinion. Acute GVHD remains a clinical diagnosis; however, GVHD is considered present if it is diagnosed and treated despite a nonconfirmatory biopsy. Additionally, we developed guidance for GVHD staging when symptoms are present in multiple target organs but only 1 has been biopsied. To summarize, biopsy results from 1 organ cannot confirm the presence of GVHD in another symptomatic organ, but the biopsy interpretation in 1 organ may increase the confidence level for the presence of GVHD in another organ. For example, if a patient has mild skin rash and persistent nausea that have not been clinically diagnosed as GVHD and a subsequent skin biopsy confirms a diagnosis of skin GVHD, the confidence that the persistent nausea is due to upper GI GVHD increases from *possible* to *probable*. Additionally, GVHD symptoms may resolve in 1 target organ after initiation of systemic therapy after biopsy for a separate organ, and we consider GVHD to be probable in the non-biopsied site unless there is compelling evidence that an etiology other than GVHD is present (eg, documented enteric infection).

MULTIPLE CONFIRMED ETIOLOGIES

We stage target organs solely based on symptom quantification. Some transplant centers will downgrade GVHD clinical severity when, in addition to GVHD, another etiology of the GVHD symptoms is documented (eg, infectious enteritis or hyperbilirubinemia due to sinusoidal obstruction syndrome). We have elected not to implement this staging rule within our consortium because it is not a standard practice at most participating centers. We do collect the raw symptom data for every target organ and retain the capability to restage GVHD if we choose to adopt this practice or a new clinical staging system altogether in the future.

CLINICAL VIGNETTES

The following clinical vignettes have been selected to illustrate the application of our clinical research data collection guidelines. All cases are actual patients treated at institutions participating in our consortium that were discussed during international adjudication webinars.

Clinical Case 1

A 37-year-old woman presents to the clinic on post-transplant day 35 after an 8/8 HLA-matched unrelated donor BMT with complaints of up to 4 watery stools daily. She denies upper GI symptoms, her liver function tests are within normal limits, and she has no rash. She is admitted to the hospital for further evaluation, including stool studies for infectious etiologies and sigmoidoscopy with biopsies.

On the day of admission, the patient's maximal stool output is estimated as 800 mL/day (4 stools estimated at 200 mL each). Because no treatment is started and there is no biopsy information available, the lower GI GVHD symptoms are assigned a confidence level of *possible*.

Sigmoid colon biopsies are obtained in post-BMT day 36. The patient has 1300 mL of diarrhea on day 36. On day 37 the biopsy report states, "Scattered single-cell necrosis and apoptotic debris, consistent with grade I graft-versus-host disease." Methylprednisolone is started at a dose of 2 mg/kg for GVHD.

In this setting the presence of lower GI GVHD is *confirmed* by a *positive* biopsy on post-BMT day 36. The highest stool output leading up to the diagnosis of GVHD is 1300 mL/day; thus, lower GI GVHD was documented as stage 2 (Table 1). Of note, the pathologist provided a histologic grade for the severity of GVHD seen on biopsy: Histopathologic grading does not correlate with clinical symptom severity and does not influence the reported clinical staging [20]. Although symptoms were first documented on day 35, the clinician did not elect to stage or treat the lower GI symptoms as GVHD before ruling out infectious causes or obtaining biopsies of the lower GI tract. GVHD onset was thus recorded as day 36 (the date of the positive biopsy) with staging as follows: skin 0 (negative), liver 0 (negative), upper GI 0 (negative), lower GI 2 (confirmed), overall grade III.

Clinical Case 2

A 67-year-old man receives a 10/10 HLA-matched peripheral blood stem cell transplant after myeloablative conditioning consisting of thiotepa, busulfan, fludarabine, and antithymocyte globulin for primary myelofibrosis. He is receiving cyclosporine and mycophenolate mofetil for GVHD prophylaxis. He develops hyperbilirubinemia during the conditioning regimen that persists through the early transplant course, ranging from 3.9 to 6.6 md/dL, which is attributed to conditioning regimen—related hepatotoxicity. On post-transplant day 10 he develops fever accompanied by new-onset diarrhea of 2000 mL/day and an erythematous macular rash involving his face and upper torso, determined to be 11% BSA. His bilirubin is 4.9 on that day. The treating clinician initiates steroids at a dose of 2 mg/kg for presumed GVHD.

Because systemic treatment was initiated for presumed GVHD, day 10 is considered GVHD onset. Skin and lower GI staging at onset are 1 (11% BSA) and 3 (2000 mL/day), respectively, and assigned a *probable* confidence level. In this scenario the hyperbilirubinemia was present before the clinical diagnosis of GVHD without a significant increase in bilirubin levels at the onset of new GVHD symptoms, so the liver was staged as 0 with a *negative* confidence level and attributed to conditioning regimen toxicity. As with signs and symptoms present in any target organ, the total bilirubin level is still collected for research purposes. Thus, staging was skin 1 (probable), liver 0 (negative), upper GI 0 (negative), lower GI 3 (probable), overall grade III.

On post-transplant day 15 the rash resolved but the diarrhea persisted, prompting an esophagogastroduodenoscopy to obtain biopsies reported as "histopathologic findings are limited and nonspecific; no GVHD or CMV is found." The 3-day average stool volume is 800 mL/day and total bilirubin level is 2.5 mg/dL and down-trending. The clinician continues to stage and treat the patient's diarrhea as GVHD.

The biopsy results are interpreted as *nondiagnostic*, despite the ongoing clinical symptoms; thus, the confidence levels remain *probable* for any active GVHD. Staging is reported as skin 0 (probable), liver 0 (negative), upper GI 0 (negative), lower GI 1 (probable), overall grade II.

Clinical Case 3

A 43-year-old woman presents to the outpatient clinic 52 days after a matched related donor BMT with a new erythematous maculopapular rash over her face, upper chest, and forearms after working in her garden. She has no other symptoms of concern for GVHD. The clinician estimates the rash involves 20% BSA and believes the rash is secondary to sun sensitivity related to voriconazole fungal prophylaxis but documents a lesser concern for new-onset acute GVHD. He

discontinues her voriconazole and treats the rash with topical steroid creams with close follow-up.

At this time the clinician favors a skin photosensitivity reaction and does not make a clinical diagnosis of GVHD. The skin rash is reported as *possible* GVHD but is not considered GVHD onset because an alternative diagnosis is favored.

Upon follow-up on post-transplant day 56, the rash has now spread to her shoulders and upper back and now involves 35% BSA despite the previous interventions. She has no other concerning findings. The clinician does not obtain a skin biopsy but now believes the rash is due to skin GVHD and elects to continue treatment with topical steroids.

Day 56 is considered GVHD onset, with staging reported as skin 2 (probable), liver 0 (negative), upper GI 0 (negative), lower GI 0 (negative), overall grade I. Although skin symptoms began on day 52, it is possible that the patient's rash was a photosensitivity reaction that later evolved into skin GVHD, and one cannot presume that GVHD was initially present.

The patient returns to clinic on post-transplant day 62, and the rash now involves the entire torso, face, bilateral forearms, and bilateral posterior thighs, reported as 60% BSA. She does not have any other symptoms concerning for GVHD. The clinician initiates systemic corticosteroids at a dose of 2 mg/kg.

The progression of the rash and initiation of systemic corticosteroid therapy does not change the confidence level for the skin GVHD because there is no confirmatory biopsy. The patient's GVHD is thus staged on day 62 as skin 3 (probable), liver 0 (negative), upper GI 0 (negative), lower GI 0 (negative), overall grade II.

DISCUSSION

Clinical data capture can prove extremely challenging in the setting of allogeneic BMT given the multitude of complications patients may experience. Our consortium believes the guidance outlined in this report is a starting point for standardizing data reporting across centers with varying clinical practices and provides structure for previously unaddressed issues for GVHD staging. It creates a systematic approach for the determination of GVHD onset, confidence in the attribution of symptoms to acute GVHD, and quantification of clinical severity of GVHD in each target organ. These guidelines also provide structure and categorization to the interpretation of biopsy reports and how the influence of a biopsy from 1 target organ guides confidence in the diagnosis of GVHD in other symptomatic organs. Some of the particularly challenging areas of GVHD clinical reporting, such as determination of the presence of upper GI GVHD without confirmatory biopsy or whether to document and stage transaminitis as liver GVHD, could be further clarified through the use of diagnostic biomarkers. Unfortunately, although validated plasma biomarkers specific to skin and lower GI GVHD have been identified [23,24], GVHD biomarkers specific to the liver and upper GI tract have yet to be identified.

Although some of the rules implemented by our consortium are arbitrary, they reflect our consensus opinion and are applied consistently across all participating centers. The intent is to standardize data extraction from the medical record without restricting autonomy of the treating clinician. We acknowledge there are limitations in our staging guidance and it does not account for every conceivable scenario. The guidance will be further refined as we encounter complex issues not currently addressed. Confidence levels

provide additional granularity to the attribution of symptoms concerning for GVHD and may help with the retrospective selection (or exclusion) of patients for research and reports and for the selection of patients for clinical trials of GVHD therapy. Future research, by our consortium and others, is required to determine the reliability of this clinical data reporting tool. For example, our GVHD onset rules can be compared with validated GVHD biomarkers in blood samples obtained at the same time. Likewise, we can test whether GVHD outcomes can be better predicted when GVHD clinical severity is incorporated into a GVHD biomarker algorithm [10].

Treatment of GVHD has remain unchanged for over 40 years; no therapies have been identified that improve outcomes when supplementing or replacing systemic corticosteroid treatment, and no treatment has been shown to be reliably effective for steroid-refractory GVHD when tested in phase III clinical trials [23-27]. Institutional biases and longstanding practices are difficult to overcome and may explain why the findings of single-center studies of GVHD therapies generally have not reproduced in the multicenter setting. Uniform clinical GVHD data reporting among centers relies on acceptance and consistent application of clear guidance and diligence in documentation and is a necessary process if we are to improve future multicenter GVHD studies. The guidance we report here has been in use internationally for 2 years and can provide much needed standardization in clinical staging for future research endeavors.

REFERENCES

1. Weisdorf DJ, Hurd D, Carter S, et al. Prospective grading of graft-versus-host disease after unrelated donor marrow transplantation: a grading algorithm versus blinded expert panel review. *Biol Blood Marrow Transplant.* 2003;9:512-518.
2. Atkinson K, Horowitz MM, Biggs JC, et al. The clinical diagnosis of acute graft-versus-host disease: a diversity of views amongst marrow transplant centers. *Bone Marrow Transplant.* 1988;3:5-10.
3. MacMillan ML, Weisdorf DJ, Wagner JE, et al. Response of 443 patients to steroids as primary therapy for acute graft-versus-host disease: comparison of grading systems. *Biol Blood Marrow Transplant.* 2002;8: 387-394.
4. Gooley TA, Chien JW, Pergam SA, et al. Reduced mortality after allogeneic hematopoietic-cell transplantation. *N Engl J Med.* 2010;363: 2091-2101.
5. Levine JE, Hogan WJ, Harris AC, et al. Improved accuracy of acute graft-versus-host disease staging among multiple centers. *Best Pract Res Clin Haematol.* 2014;27:283-287.
6. Przepiorka D, Weisdorf D, Martin P, et al. 1994 Consensus Conference on Acute GVHD Grading. *Bone Marrow Transplant.* 1995;15:825-828.
7. MacMillan ML, Robin M, Harris AC, 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;21:761-767.
8. Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. *Lancet.* 2009;373:1550-1561.
9. Wallace AB. The exposure treatment of burns. *Lancet.* 1951;1:501-504.
10. Levine JE, Braun TM, Harris AC, et al. A prognostic score for acute graft-versus-host disease based on biomarkers: a multicenter study. *Lancet Haematol.* 2015;2:e21-e29.
11. Leisenring WM, Martin PJ, Petersdorf EW, et al. An acute graft-versus-host disease activity index to predict survival after hematopoietic cell transplantation with myeloablative conditioning regimens. *Blood.* 2006;108:749-755.

12. Robin M, Porcher R, de Castro R, et al. Initial liver involvement in acute GVHD is predictive for nonrelapse mortality. *Transplantation*. 2009;88: 1131-1136.
13. Glucksberg H, Storb R, Fefer A, et al. Clinical manifestations of graft-versus-host disease in human recipients of marrow from HL-A-matched sibling donors. *Transplantation*. 1974;18:295-304.
14. Thomas ED, Storb R, Clift CD, et al. Bone-marrow transplantation. *N Engl J Med*. 1975;292(Pt. 2):895-902.
15. Martin PJ, McDonald GB, Sanders JE, et al. Increasingly frequent diagnosis of acute gastrointestinal graft-versus-host disease after allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant*. 2004;10:320-327.
16. Weisdorf D, Snover D, Haake R, et al. Acute upper gastrointestinal graft-versus-host disease: clinical significance and response to immunosuppressive therapy. *Blood*. 1990;76:624-629.
17. Castilla-Llorente C, Martin PJ, McDonald GB, et al. Prognostic factors and outcomes of severe gastrointestinal GVHD after allogeneic hematopoietic cell transplantation. *Bone Marrow Transplant*. 2014;49:966-971.
18. Abraham J, Janin A, Gornet JM, et al. Clinical severity scores in gastrointestinal graft-versus-host disease. *Transplantation*. 2014;97:965-971.
19. Nevo S, Enger C, Swan V, et al. Acute bleeding after allogeneic bone marrow transplantation: association with graft versus host disease and effect on survival. *Transplantation*. 1999;67:681-689.
20. Shulman HM, Kleiner D, Lee SJ, et al. Histopathologic diagnosis of chronic graft-versus-host disease: National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease. II. Pathology Working Group Report. *Biol Blood Marrow Transplant*. 2006;12:31-47.
21. Kuykendall TD, Smoller BR. Lack of specificity in skin biopsy specimens to assess for acute graft-versus-host disease in initial 3 weeks after bone-marrow transplantation. *J Am Acad Dermatol*. 2003;49: 1081-1085.
22. Massi D, Fondi C, Nozzoli C, et al. The impact of histopathologic examination of graft-versus-host disease in the era of reduced-intensity conditioning regimen: a study from the Gruppo Italiano Trapianto di Midollo Osseo. *Hum Pathol*. 2011;42:254-268.
23. Ho VT, Cutler C. Current and novel therapies in acute GVHD. *Best Pract Res Clin Haematol*. 2008;21:223-237.
24. Martin PJ, Inamoto Y, Flowers ME, Carpenter PA. Secondary treatment of acute graft-versus-host disease: a critical review. *Biol Blood Marrow Transplant*. 2012;18:982-988.
25. Xhaard A, Rocha V, Bueno B, et al. Steroid-refractory acute GVHD: lack of long-term improved survival using new generation anticytokine treatment. *Biol Blood Marrow Transplant*. 2012;18:406-413.
26. Ruutu T, Gratwohl A, de Witte T, et al. Prophylaxis and treatment of GVHD: EBMT-ELN working group recommendations for a standardized practice. *Bone Marrow Transplant*. 2014;49:168-173.
27. Bolanos-Meade J, Logan BR, Alousi AM, et al. Phase 3 clinical trial of steroids/mycophenolate mofetil vs steroids/placebo as therapy for acute GVHD: BMT CTN 0802. *Blood*. 2014;124:3221-3227.