ASBMT American Society for Blood and Marrow Transplantation

Diagnosis and Staging of Chronic Graft-versus-Host Disease in the Clinical Practice

Hildegard T. Greinix,[1] Christoph Loddenkemper,[2,3] Steven Z. Pavletic,[4] Ernst Holler,[5] Gerard Socié,[6] Anita Lawitschka,[7] Joerg Halter,[8] Daniel Wolff[5]

Based on expert opinion and retrospective data the National Institutes of Health (NIH) Consensus Development Project proposed criteria for diagnosis and staging of both overall severity as well as organ severity of chronic graft-versus-host disease (cGVHD) for use in clinical trials. In 2008, representatives of German and Austrian allogeneic hematopoietic stem cell transplant (HSCT) centers established a study group on cGVHD during the annual meeting of the German Working Group on Bone Marrow and Blood Stem Cell Transplantation (DAG-KBT) to intensify a dialog among HSCT physicians, pathologists, and medical consultants focusing on the usefulness of the NIH consensus criteria for patient care in clinical practice and to promote collaborations between HSCT centers as well as different medical specialities involved in HSCT. We first conducted a survey of current practices of diagnosis, staging, and overall grading of cGVHD in daily clinical routine by sending an electronic questionnaire to the heads of the HSCT centers. During 3 meetings in 2009, more representatives of allogeneic HSCT centers were included into the discussion process, resulting in 81% participation representing 88% of all allogeneic HSCT activities in Germany, Austria, and Switzerland. During the third consensus meeting held in Regensburg, Germany, from November 6 to November 7, 2009, important agreements were achieved among participant having a strong impact on care of patients with cGVHD. Areas of disagreement such as distinction between classical NIH cGVHD and overlap syndrome or assignment of liver GVHD after day 100 to acute or chronic category will be further assessed in prospective observational studies among participants in the near future.

Biol Blood Marrow Transplant 17: 167-175 (2011) © 2011 American Society for Blood and Marrow Transplantation

KEY WORDS: Chronic GVHD, Staging, Severity scoring

INTRODUCTION

Chronic graft-versus-host disease (cGVHD) is a serious complication of allogeneic hematopoietic stem cell transplantation (HSCT) and has a significant negative impact on patient survival and quality of life [1-3]. It is a multisystemic disorder and can present with a variety of clinical signs and symptoms resembling autoimmune diseases such as scleroderma, Sjogren's syndrome, bronchiolitis obliterans, and chronic immunodeficiency. The original descriptions of cGVHD and its staging of severity were based on small numbers of patients in the precyclosporine era [4]. Recently, the National Institutes of Health (NIH) consensus development project on criteria for clinical trials in cGVHD established standardized criteria for the diagnosis of cGVHD and proposed tools for scoring cGVHD organ involvement and overall severity [5]. The recommendations of the NIH working group represent a consensus opinion based on leading international HSCT expert assessments and evaluation of peer-reviewed literature. The proposed methods and tools for diagnosis and scoring of cGVHD, however, have to be validated in prospective studies. Furthermore, their use in daily clinical routine has not been evaluated yet with regard to their feasibility and accuracy because they were established for use in clinical trials. In large HSCT outpatient clinics routine assessment of cGVHD patients can be time-consuming and occupies valuable and costly resources. cGVHD is a complex medical

From the [1]Department of Internal Medicine I, Bone Marrow Transplantation, Medical University of Vienna, Vienna, Austria; [2]Institute of Pathology/Research Center ImmunoSciences (RCIS), Charite'–Universitaetsmedizin Berlin, Berlin, Germany; [3]Institute of Pathology, Technische Universitaet Muenchen, Munich, Germany; [4]Experimental Transplantation and Immunology Branch, National Cancer Institute, NIH, Bethesda, Maryland; [5]Department of Hematology and Oncology, University of Regensburg, Germany; [6]Service d' Hematologie Greffe & Inserm Hopital Saint-Louis, Paris, France; [7]Department of Stem Cell Transplantation, St. Anna Children's Hospital, Vienna, Austria; and [8]Department of Hematology, University hospital of Basel, Basel, Switzerland.

Financial disclosure: See Acknowledgments on page 173.

Correspondence and reprint requests: Hildegard T. Greinix, MD, Medizinische Universitaet Wien, Klinik fuer Innere Medizin I, Knochenmarktransplantation Waehringer Guertel 18-20, A-1090 Vienna, Austria (e-mail: hildegard.greinix@meduniwien.ac.at).

Received February 9, 2010; accepted July 19, 2010

1083-8791/$36.00
doi:10.1016/j.bbmt.2010.07.017

condition, and patients' care requires multidisciplinary teams of physicians, nurses, pharmacists, psychosocial personnel, and other clinical staff with specialized diagnostic and treatment units. Because patients with cGVHD are at high risk for infections, organ toxicity, and other life-threatening treatment-related complications for prolonged periods of time, a substantial number of prevention and monitoring strategies and supportive care practices are already common and well established in most HSCT units. Therefore, additional diagnostic procedures, for example, for serial monitoring of cGVHD patients, could possibly put an enormous additional burden on outpatient clinics and economic constraints have to be considered.

In 2008, representatives of German and Austrian allogeneic HSCT centers (see Appendix) established a study group on cGVHD during the annual meeting of the German Working Group on Bone Marrow and Blood Stem Cell Transplantation (DAG-KBT) and first conducted a survey of current practices of diagnosis, staging, and overall grading of cGVHD in daily clinical routine by sending an electronic questionnaire regarding the use of NIH consensus definitions of acute GVHD (aGVHD) and cGVHD, use of NIH consensus scoring of organ-specific manifestations and their severity, and prognostic criteria in daily clinical practice to the heads of the HSCT centers. During 2009, the study group on cGVHD expanded from the initial 21 HSCT centers' representatives to eventually 58 (81%) representatives of allogeneic HSCT centers including also Switzerland. The results of the survey as well as the published NIH consensus documents were discussed in 3 meetings with regard to acceptance or rejection for daily clinical routine and in the final meeting held in Regensburg, Germany, from November 6 to November 7, 2009, a consensus document was finalized. The conferences were organized under the auspices of the DAG-KBT and the German Society of Hematology and Oncology (DGHO), the Austrian Stem Cell Transplant Working Group, and the Austrian Society of Hematology and Oncology (OEGHO), the Swiss Blood Stem Cell Transplantation Group (SBST), and the German-Austrian Paediatric Working Group on Stem Cell Transplantation (PÄD-AG-KBT). Of note, pediatric aspects of the consensus document were also discussed within a meeting of the PÄD-AG-KBT in fall 2009 and were accepted for daily clinical practice.

Results from these activities allow comment on areas of agreement and controversy and identify areas for future work.

DESCRIPTION OF THE GERMAN/AUSTRIAN/ SWISS STUDY GROUP ON cGVHD

In the fall of 2008, physicians from the DAG-KBT at an attending level position representing 21 allogeneic HSCT centers met in Wiesbaden, Germany, during their annual meeting to discuss the recently published NIH consensus criteria on cGVHD [5-8] with regard to their feasibility in daily clinical patient care. Participants decided to initiate a consensus consortium on cGVHD with study groups on diagnosis and staging, first-line therapy, salvage therapy, and supportive care. To evaluate the current clinical practice a questionnaire on diagnosis and staging was developed by the study group that was sent electronically to the medical directors of these 21 allogeneic HSCT centers. Within 4 weeks 13 of 21 (62%) responded and center characteristics are shown in Table 1a. In further meetings of the consensus consortium in March, June, and November 2009, the number of participating centers rose to finally 52 of 62 German (84%), 5/6 (83%) Austrian, and 1 of 4 (25%) Swiss centers performing 88% of all allogeneic HSCT of their countries. Characteristics of all responding and nonresponding centers are shown in Table 1b. In addition, the consensus on diagnosis and staging of cGVHD was discussed within the German-Austrian Paediatric Working Group on Stem Cell Transplantation in the fall of 2009.

RESULTS OF THE INITIAL SURVEY ON DAILY CLINICAL PRACTICE OF ALLOGENEIC HSCT CENTERS

As shown in Table 2, 11 of 13 (85%) centers stated to use the NIH criteria for diagnosis of cGVHD in daily clinical routine. Ten (77%) also distinguished between classic NIH cGVHD and overlap syndrome. All agreed with the NIH definition of classic cGVHD and with the definitions of diagnostic and distinctive signs of cutaneous cGVHD. Of note, only 5 (38%) routinely performed skin biopsies for confirmation of cGVHD. Whereas acceptance of NIH definitions on cGVHD of muscles, fascia, and joints was high among centers, only 1 (8%) stated to routinely perform an electromyography in case of muscle cramps. Acceptance of severity definitions both of organ manifestations as well as overall severity of cGVHD among centers was high as shown in Table 2. Only the definition of severity of GI involvement had lower acceptance.

In addition, the survey asked about features of cGVHD that are known risk factors for treatment-related mortality (TRM), revealing a high acceptance of thrombocytopenia at onset of cGVHD and progressive onset of cGVHD as risk factors for patient outcome.

AGREEMENT/DISAGREEMENT OF PARTICIPANTS OF CONSENSUS MEETINGS WITH NIH DEFINITION OF OVERLAP SYNDROME

Currently, it is unclear whether pathophysiologic mechanisms involved in classical cGVHD differ from

Table 1a. Characteristics of Centers Participating in Initial Survey (n = 21)

	Responding Centers N (%)	Nonresponding Centers N (%)	All Centers N
Number of centers	13 (62)	8 (38)	21
Number of centers with adult patients	12 (63)	7 (37)	19
Number of centers with pediatric patients	1 (50)	1 (50)	2
Number of alloHSCT per year	959 (68)	456 (32)	1415
Number of alloHSCT with matched related donor per year	257 (68)	119 (32)	376
Number of alloHSCT with unrelated donor per year	670 (68)	314 (32)	984
Number of alloHSCT with cord blood per year	7 (70)	3 (30)	10
Number of alloHSCT with mismatched related donor per year	25 (56)	20(44)	45

N indicates number; alloHSCT, allogeneic hematopoietic stem cell transplantation.

the ones leading to overlap syndrome, and thus, about one-third of participants in the consensus meetings felt no need for distinction between these 2 subtypes in daily clinical practice. In the discussions, the current available evidence on prognosis of patients with overlap syndrome was presented. Jagasia and colleagues [9] retrospectively reclassified 110 patients and reported a significantly worse survival of patients with any features of aGVHD after day 100 of HSCT compared with NIH cGVHD. In a retrospective analysis of 54 consecutive patients, Arora and colleagues [10] observed a significantly worse survival at 3 years in patients with late, persistent, or recurrent aGVHD of 25% compared with 87% in patients with overlap syndrome and 75% in patients with classic NIH cGVHD. Recently, Cho and colleagues [11] retrospectively reclassified 211 patients who developed GVHD more than 100 days after HSCT and observed 21% late aGVHD, 30% overlap syndrome, and 49% classic NIH cGVHD. Four-year GVHD-specific survival was significantly different among patients with late-onset, persistent, and recurrent aGVHD (100% versus 86% versus 56%, $P = .009$), but did not differ significantly between overlap syndrome and classic NIH cGVHD. Similar findings were reported by Vigorito and colleagues [12] in 740 patients. Considering that about 15% of patients with historic cGVHD may be misdiagnosed and have late-onset, persistent, or recurrent ("late") aGVHD, and that these patients have significantly worse survival rates, a distinction between late aGVHD and overlap syndrome seems very important. Whether the outcome of patients with overlap syndrome differs from the ones with classic NIH cGVHD should be assessed in prospective studies. Therefore, all participants of the German/Austrian/Swiss consensus meetings eventually agreed on documenting the category "overlap syndrome" separately from classic NIH cGVHD, which is in agreement with others suggesting proper stratification of patients according to the NIH GVHD categories for their inclusion in clinical trials [13].

AGREEMENT/DISAGREEMENT OF PARTICIPANTS OF CONSENSUS MEETINGS WITH NIH DEFINITION OF LIVER MANIFESTATIONS OF cGVHD

Hepatic cGVHD can present as cholestasis with increased bilirubin or alkaline phosphatase but also as acute hepatitis with elevation of liver enzymes [14,15]. According to the NIH Consensus Development Project liver biopsy is required to confirm GVHD involvement of the liver [5]. In the survey, only around half of all centers stated to perform liver biopsies in case of isolated elevations of liver enzymes as shown in Table 2. During the consensus meetings, various clinical scenarios were discussed, and mainly lack of response to immunosuppressive therapies and possible concomitant infection were mentioned as indications for liver biopsies in daily clinical routine. Because, according to the NIH consensus, histologic similarities between acute and chronic liver GVHD do not allow to have a definitive diagnosis of chronic liver GVHD when no other organ has a diagnostic or distinctive sign of cGVHD [5,6], isolated elevations of liver function tests remain challenging in daily clinical practice. In the discussion among participants of the German/Austrian/Swiss consensus meetings, it was

Table 1b. Characteristics of Centers Participating in Consensus Process

	Participating Centers N (%)	Nonparticipating Centers N (%)	All Centers N
Number of centers	58 (81)	14 (19)	72
Number of centers with adult patients	35 (74)	12* (26)	47*
Number of centers with pediatric patients	23 (88)	3 (12)	26
Number of alloHSCT per year	2383 (88)	331 (12)	2714
Number of alloHSCT with matched related donor per year	655 (86)	103 (14)	758
Number of alloHSCT with unrelated donor per year	1583 (89)	204 (11)	1787
Number of alloHSCT with cord blood per year	28 (87.5)	4 (12.5)	32
Number of alloHSCT with mismatched related donor per year	117 (85)	20 (15)	137

N indicates number; alloHSCT, allogeneic hematopoietic stem cell transplantation.
*In Innsbruck, only the pediatric HSCT program participated in the consensus process.

Table 2. Results of Survey on Daily Practice of Diagnosis and Staging of Chronic GVHD

Question	No Centers in Agreement (%)
Do you use the NIH criteria for diagnosis of cGVHD in daily routine?	11 (85)
Do you distinguish between classic cGVHD and overlap syndrome?	10 (77)
Do you agree with the definition of classic cGVHD?	13 (100)
Do you agree with the definition of overlap syndrome?	12 (92)
At least 1 diagnostic sign* is necessary for diagnosis of cGVHD.	13 (100)
When a diagnostic sign* is missing, a distinctive sign with confirmation by lab, radiology, or biopsy is required for diagnosis of cGVHD.	12 (92)
Do you agree with the diagnostic signs* of cGVHD of the skin?	13 (100)
Do you agree with the distinctive signs* of cGVHD of the skin?	13 (100)
Do you document the % of superficial sclerosis of the skin?	8 (61.5)
Do you document the % of deep sclerosis of the skin?	6 (46)
Do you document the % of hypo/hyperpigmentation of the skin?	3 (23)
We perform routinely skin biopsies in cGVHD patients.	5 (38)
Do you agree with the diagnostic signs* of cGVHD of the oral mucosa?	12 (92)
Do you agree with the distinctive signs* of cGVHD of the oral mucosa?	11 (85)
We routinely exclude infections of the oral mucosa.	9 (69)
Do you agree with the distinctive signs* of cGVHD of the eyes?	12 (92)
We routinely perform Schirmer tests.	4 (31)
Symptomatic patients are routinely seen by an ophthalmologist.	13 (100)
Do you agree with the diagnostic signs* of cGVHD of the genitalia?	11 (85)
Do you agree with the distinctive signs* of cGVHD of the genitalia?	11 (85)
Symptomatic patients are routinely seen by a gynecologist.	12 (92)
We routinely exclude infections of the genitalia.	3 (23)
Do you agree with the diagnostic signs* of cGVHD of the GI tract?	12 (92)
In symptomatic patients endoscopies are performed routinely.	11 (85)
In symptomatic patients endoscopic biopsies are performed routinely.	11 (85)
In case of weight loss resorption tests are performed routinely.	1 (8)
In case of diarrhea stool cultures are performed routinely.	13 (100)
In case of isolated elevation of liver function tests liver biopsies are done.	8 (61.5)
The diagnosis of cGVHD of the liver requires a distinctive sign in at least 1 other organ system.	6 (46)
We routinely exclude viral infections in case of elevated liver function tests.	13 (100)
In case of elevated liver function tests ultrasound/CT scan of the liver is routinely performed.	12 (92)
Do you agree with the diagnostic signs* of cGVHD of the lung?	9 (69)
Do you agree with the distinctive signs* of cGVHD of the lung?	10 (77)
We routinely perform pulmonary function tests after HSCT.	9 (69)
In case of decline of FEV_1 an HR-CT scan is routinely performed.	12 (92)
In case of decline of FEV_1 a BAL is routinely performed.	4 (31)
In case of possible BO lung biopsy is performed.	8 (61.5)
Do you agree with the diagnostic signs* of cGVHD of muscles/fascia/joints?	13 (100)
Do you agree with the distinctive signs* of cGVHD of muscles/fascia/joints?	11 (85)
In case of muscle cramps EMG is routinely performed.	1 (8)
In cGVHD pts tests for autoantibodies are routinely done.	7 (54)
Do you agree with the definition* of severity of cGVHD of the skin?	11 (85)
Do you agree with the definition* of severity of cGVHD of the oral mucosa?	12 (92)
Do you agree with the definition* of severity of cGVHD of the eyes?	13 (100)
Do you agree with the definition* of severity of cGVHD of the GI tract?	10 (77)
Do you agree with the definition* of severity of cGVHD of the liver?	12 (92)
Do you agree with the definition* of severity of cGVHD of the lung?	11 (85)
Do you agree with the definition* of severity of cGVHD of the joints?	13 (100)
Do you agree with the definition* of severity of cGVHD of the genitalia?	13 (100)
Do you agree with the definition* of mild cGVHD?	13 (100)
Do you agree with the definition* of moderate cGVHD?	12 (92)
Do you agree with the definition* of severe cGVHD?	13 (100)
Thrombocytopenia at onset of cGVHD is a risk factor for TRM.	11 (85)
Progressive onset of cGVHD is a risk factor for TRM.	12 (92)
>50% skin involvement by cGVHD is a risk factor for TRM.	10 (77)

No indicates number; HR-CT, high-resolution chest computed tomography; BAL, bronchoalveolar lavage; GI, gastrointestinal; EMG, electromyography; TRM, treatment-related mortality; cGVHD, chronic graft-versus-host disease; HSCT, hematopoietic stem cell transplant; BO, bronchiolitis obliterans.
*According to [5].

suggested to document liver involvement according to the categories "Proven," "Probable," and "Possible" when isolated elevations of liver function tests such as serum bilirubin or gamma GT are present in a patient. Proven liver involvement could be defined as histologic evidence of portal fibrosis, marked loss of bile ducts, and chronic cholestasis with bile ductular proliferation with or without bridging fibrosis that reflects chronicity [6]. Probable liver involvement can be considered in case of cholestasis with increased bilirubin or alkaline phosphatase and onset in close proximity to discontinuation of immunosuppression and

lack of concomitant infection or medication with documented liver toxicity. Possible liver involvement could be defined as presentation of elevated liver enzymes in close proximity to discontinuation of immunosuppression or donor lymphocyte infusions but indicates alternate diagnoses and reasons for suspicion. Others suggested to document liver GVHD without assignment to the acute or chronic category but with further information such as time of onset of liver GVHD after donor lymphocyte infusions or discontinuation of immunosuppression. This would allow an assignment retrospectively when the patient has a longer follow-up and the course of GVHD becomes more conclusive.

In view of the difficulties of distinguishing GVHD from other diseases such as drug toxicities or infections and the heterogeneous histologic presentations, a close cooperation with a pathologist knowledgeable in liver diseases including GVHD would improve our diagnostic arsenal substantially. It was agreed by all participants that a cooperation on a national level with reference pathologists should be established to improve diagnosis of liver GVHD. A main indication for liver biopsies is clinical refractoriness to immunosuppressive therapy allowing immunohistologic assessment of extent of inflammation and bile duct damage. Experience-based observations indicate that the time to recovery after immunosuppressive treatment is proportional to the degree of ductopenia [6] and thus, immunohistologic results can support response assessment. Besides verifying the diagnosis of liver GVHD, other processes such as viral infections or hemosiderosis can be excluded histopathologically.

AGREEMENT/DISAGREEMENT OF PARTICIPANTS OF CONSENSUS MEETINGS WITH NIH DEFINITION OF LUNG MANIFESTATIONS OF cGVHD

According to the NIH Consensus the only diagnostic manifestation of lung cGVHD is biopsy-proven bronchiolitis obliterans (BO) [5]. BO diagnosed via pulmonary function and radiologic testing (bronchiolitis obliterans syndrome [BOS]) requires at least 1 other distinctive manifestation in a separate organ system to establish the diagnosis of cGVHD. In the survey 9 (69%) centers stated to routinely perform pulmonary function tests after HSCT as shown in Table 2. In the case of a decrease in forced expiratory volume in 1 second (FEV_1), the vast majority of centers performs high-resolution (HR) CT scans of the lungs but only about a third stated to routinely perform a bronchoalveolar lavage (BAL). In the case of documented changes consistent with BOS in chest HR-CT scans, the vast majority of centers stated to perform a BAL but only few of them considered a transbronchial biopsy mainly to exclude infections or toxicities. Thus, during the German/Austrian/Swiss consensus meetings the discussions focused on screening of asymptomatic patients for early detection of cGVHD of the lungs because only 7 (54%) centers participating in the survey performed serial pulmonary function tests starting at various time points and as late as 12 months after HSCT, and few of them also included asymptomatic patients. Critics of this policy mentioned the additional logistical and economical burden on HSCT facilities by serial pulmonary function tests and were reluctant to see their advantage for routine patient care. In support of serial pulmonary function tests Chien and colleagues' [16] publication on an association of airflow decline by day 100 after HSCT with a significantly increased risk for development of transplant-related airflow obstruction at 1 year after HSCT in a retrospective cohort analysis of 1892 patients was presented during the consensus meetings. However, a single measurement of lung function on day 100 was not sufficient because airflow decline by day 100 was not associated with an increased mortality risk and patients with fast declines of FEV_1 during the first year after HSCT experienced the highest mortality risk supporting repeat pulmonary function tests after HSCT.

Recently, Gunn and colleagues [17] observed a significant correlation of severity of air trapping on HR-CT scans with changes in FEV_1 measurements in pulmonary function tests in BO patients after HSCT. In the prospective German multicenter validation trial on effect of cGVHD on quality of life and activity profile, occurrence of BO at grade 2 according to the NIH Consensus was significantly associated with worse quality of life and activity profile of patients [18]. In view of the high mortality rates of patients with BOS ranging from 14% to 100% [16,19-21] and the option of therapeutic intervention before irreversible pulmonary damage evolves, all participants of the German/Austrian/Swiss consensus meetings eventually agreed on performing serial pulmonary function tests starting around day 100 after HSCT and being repeated every 3 months within the first year. Furthermore, in case of a decrease in FEV_1 by 20% compared to pre-HSCT, chest HR-CT scan in expiration, and BAL should be performed to exclude infections.

Of note, a modification of the NIH criteria on lung involvement was proposed recently to diagnose patients with earlier disease for interventions [22].

AGREEMENT/DISAGREEMENT OF PARTICIPANTS OF CONSENSUS MEETINGS WITH NIH RECOMMENDATIONS ON OTHER DIAGNOSTIC PROCEDURES AND DOCUMENTATION IN DAILY CLINICAL ROUTINE

For skin scoring of cGVHD the NIH Consensus recommended documentation of percentage of affected

body surface area (BSA) as well as superficial and deep sclerosis [5]. This expert opinion consensus, however, has not been validated yet. Other skin scoring systems such as the total skin score validated by Greinix and colleagues [23] and used in a randomized multicenter phase II study in cGVHD patients [24] and the Hopkins scale reported by Jacobsohn and colleagues [25] were discussed during the consensus meetings. Whereas a minority of centers initially reported documentation of percentage of superficial and deep hidebound sclerosis as shown in Table 2, further discussions during the consensus meetings let to a high level of acceptance of the fact that proper documentation of both extent and quality of cutaneous involvement by cGVHD at baseline and prior to therapeutic changes is an important prerequisite for response evaluation during and after topical and systemic immunosuppressive therapy not only in clinical studies but also in daily clinical practice.

Acceptance of diagnostic and distinctive signs of oral mucosa involvement by cGVHD was also high as shown in Table 2. However, only 9 (69%) centers of the survey routinely excluded infections of the oral mucosa in symptomatic patients.

During the consensus meetings the validation study of Elad and colleagues [26] revealing a strong correlation between the total NIH score and both erythema and ulcerations but no correlation between the scores of the lichenoid/mucocele types and the scores of the other types of oral manifestations and the evaluation of inter-and intraobserver variability of NIH response criteria scoring scale for oral cGVHD by Treister and colleagues [27] using intraoral photographs were discussed.

Pediatricians emphasized that children rarely report dry mouth, taste alteration, and difficulties of swallowing, and thus, a reduction of oral intake or the need of increased drinking during meals could be often the only symptoms of cGVHD of the oral mucosa. All participants agreed on the importance of excluding infections of the oral cavity by appropriate swabs in daily clinical routine, and decided to incorporate these diagnostic procedures in their daily standards of care.

Although the vast majority of participants of the German/Austrian/Swiss consensus meetings agreed with the NIH definitions on eye involvement [5] very few centers routinely performed Schirmer tests in their patients as shown in Table 2. In further discussions, including ophthalmologists, the need for close cooperations with ophthalmologic specialists for both diagnostic procedures and topical therapies in daily clinical routine was emphasized by all participants. After thorough discussions the ophthalmologist community developed a standardized grading of severity and criteria for response evaluation for daily clinical routine that will be submitted by T. Dietrich and colleagues as a separate manuscript for publication.

In the survey 12 (92%), centers provided care by a specialist for patients with cGVHD of the genitalia but only 3 (23%) routinely excluded infections as shown in Table 2. During further discussions among participants of the German/Austrian/Swiss consensus meetings the importance of these diagnostic procedures was emphasized and accepted by all centers.

Whereas the vast majority of centers stated to routinely perform endoscopies and biopsies in symptomatic patients, only 1 (8%) referred patients to resorption tests in case of weight loss as shown in Table 2. During the consensus meetings currently available diagnostic procedures in cGVHD patients with weight loss and wasting syndrome were discussed and incorporation of resorption tests into daily clinical routine accepted by all participants.

During the German/Austrian/Swiss consensus meetings the NIH gobal severity staging criteria that have been preliminary validated for feasibility and survival impact in several recent retrospective studies [9,10,11,28] were discussed in detail and the vast majority of participants agreed with them.

PROGRESSIVE ONSET TYPE OF cGVHD AND TRANSPLANT OUTCOME

Both in the survey as well as during the German/Austrian/Swiss consensus meetings, the vast majority of participants agreed with factors associated with increased risk of TRM as stated by the NIH Consensus [5]. In addition, participants mentioned that the NIH expert opinion panel did not consider BOS and its impact on TRM. During the consensus meetings, participants discussed at length the differences and/or similarities between overlap syndrome and progressive onset of cGVHD with regard to impact on transplant outcomes including TRM. In the original definition of the Seattle group, cGVHD had a progressive onset if it followed aGVHD without resolution of aGVHD [29]. Several investigators demonstrated an association of progressive onset with increased nonrelapse mortality (NRM) in patients with historically defined cGVHD [29-34].

Using the NIH scoring system Perez-Simon and colleagues [28] recently reported significantly worse survival of patients with severe NIH cGVHD and progressive onset type. Vigorito and colleagues [12] retrospectively reclassified 740 patients according to the development of NIH cGVHD at the onset of historically defined cGVHD or at any time afterward. Antecedent late aGVHD was associated with increased risk of NRM among patients with NIH cGVHD. This association corresponds with previous results showing decreased survival among patients with a progressive onset of historically defined cGVHD from aGVHD. The authors speculated that antecedent aGVHD is associated with prolonged inflammatory insult to target

organs, a greater degree of accumulated tissue damage, and more profound immune dysregulation, making subsequent NIH cGVHD more difficult to control and contributing to increased risks of fatal infection or organ failure [12].

When Arora and colleagues [10] retrospectively reclassified 54 patients with cGVHD according to the NIH consensus criteria the 20 patients (37%) with progressive onset type turned out to have either late aGVHD (n = 5, 25%), classic cGVHD (n = 4, 20%), or overlap syndrome (n = 11, 55%). Patients with late aGVHD had a significantly higher TRM and significantly worse survival, whereas no significant differences between patients with classic NIH cGVHD and overlap syndrome were observed. Thus, progressive onset type is not necessarily identical with overlap syndrome and most likely includes a substantial number of patients with aGVHD as reported by several investigators [9-12,35].

Thus, all participants of the German/Austrian/Swiss Consensus conference agreed on documenting both presence of aGVHD on day 100 after HSCT as well as onset type of cGVHD to allow evaluation of prognostic criteria of historically defined cGVHD such as progressive onset type in cohorts of patients studied prospectively with current day immunosuppressive and supportive care.

CONCLUSIONS

cGVHD remains a source of significant morbidity and mortality after allogeneic HSCT. Despite the fact that the NIH consensus development project proposed criteria for diagnosis and classification of cGVHD for use specifically in clinical trials, these criteria are very feasible for use in daily clinical practice. For both definitions of cGVHD, as well as overall and organ specific severity staging, high rates of acceptance were obtained among the vast majority of allogeneic HSCT centers in Germany, Austria, and Switzerland. In addition, medical consultants from gastroenterology, pulmonary diseases, ophthalmology, gynecology, and others agreed on the proposed diagnostic procedures for establishing the diagnosis of cGVHD and the severity scoring of affected organs in HSCT clinical practice. Main topics of disagreement among participants of the consensus meetings were the importance of the distinction between classical NIH cGVHD and overlap syndrome and whether isolated GVHD of the liver after day 100 after HSCT should be considered as aGVHD or cGVHD. For both areas, participants agreed on more detailed documentation of all future patients to allow prospective multicenter studies assessing transplant outcomes. Whereas the distinction between classical cGVHD and overlap syndrome may not be clinically important during routine monitoring of patients and their response evaluation, the 2 different categories at onset of cGVHD will be evaluated prospectively in a registry study regarding prognosis of patients including NRM and survival. Purely clinical studies, however, will not be able to resolve the question of whether isolated liver abnormalities should be classified as late aGVHD or NIH cGVHD. It is possible that blood biomarkers or molecular studies of liver biopsies might yield some information in that regard.

As a next step, the consensus group will establish a cGVHD registry including staging and response evaluation according to NIH criteria as a prerequisite for observational studies within the consortium. Furthermore, a network of reference pathologists is currently being established within Germany, Austria, and Switzerland to improve the diagnostic yield of tissue biopsies obtained from patients with cGVHD and to discuss clinically related research projects on tissue specimens. Moreover, the close cooperation with the German/Austrian/Swiss consortium on cGVHD will allow prospective sample collection for candidate biomarker studies in the near future.

ACKNOWLEDGMENTS

The authors are grateful to Stephanie Lee, Mohamad Mohty, and Sharon Elad for their contributions and would like to thank all participants of the consensus meetings for their valuable input. The conference in Regensburg was supported by the Jose Carreras Foundation project "Competence centre Regensburg." The authors are grateful to the German Registry of Stem Cell Transplantation (DRST), the Austrian Registry of Stem Cell Transplantation (ASCT), and the Swiss National Stem Cell Transplant Registry for providing the information on annual HSCT numbers in these countries.

Financial disclosure: H. T. was supported by European Commission Grant 037703 STEMDIAGNOSTICS.

REFERENCES

1. Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2003;9:215-233.
2. Socie G, Stone JV, Wingard JR, 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.
3. Baird K, Pavletic SZ. Chronic graft versus host disease. *Curr Opin Hematol*. 2006;13:426-435.
4. Shulman HM, Sullivan KM, Weiden PL, et al. Chronic graft vs. host syndrome in man. A long-term clinicopathologic study of 20 Seattle patients. *Am J Med*. 1980;69:204-217.
5. Filipovich AH, Weisdorf D, Pavletic S, et al. National Institutes of Health Consensus Development Project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. *Biol Blood Marrow Transplant*. 2005;11:945-955.
6. 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.

7. 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.
8. Couriel D, Carpenter PA, Cutler C, et al. Ancillary therapy and supportive care of chronic graft-versus-host disease: National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: V. Ancillary therapy and supportive care working group report. *Biol Blood Marrow Transplant*. 2006;12:375-396.
9. Jagasia M, Giglia J, Chinratanalab W, et al. Incidence and outcome of chronic graft-versus-host disease using National Institutes of Health Consensus criteria. *Biol Blood Marrow Transplant*. 2007; 13:1207-1215.
10. Arora M, Nagaraj S, Witte J, et al. New classification of chronic GVHD: added clarity from the consensus diagnoses. *Bone Marrow Transplant*. 2009;43:149-153.
11. Cho BS, Min CK, Eom KS. Feasibility of NIH consensus criteria for chronic graft-versus-host disease. *Leukemia*. 2009;23:78-84.
12. Vigorito AC, Campregher PV, Storer BE, et al. Evaluation of NIH consensus criteria for classification of late acute and chronic GVHD. *Blood*. 2009;114:702-708.
13. Pavletic SZ, Lee SJ, Socie G, Vogelsang G. Chronic graft-versus-host disease: implications of the National Institutes of Health consensus development project on criteria for clinical trials. *Bone Marrow Transplant*. 2006;38:645-651.
14. Shulman HM, Sharma P, Amos D, et al. A coded histologic study of hepatic graft-versus-host disease after human bone marrow transplantation. *Hepatology*. 1988;8:463-470.
15. Strasser SI, Shulman HM, Flowers ME, et al. Chronic graft-vs-host disease of the liver: presentation as an acute hepatitis. *Hepatology*. 2000;32:1265-1271.
16. Chien JW, Martin PJ, Flowers ME, Nichols WG, Clark JG. Implications of early airflow decline after myeloablative allogeneic stem cell transplantation. *Bone Marrow Transplant*. 2004;33: 759-764.
17. Gunn ML, Godwin JD, Kanne JP, Flowers ME, Chien JW. High-resolution CT findings of bronchiolitis obliterans syndrome after hematopoietic stem cell transplantation. *J Thorac Imaging*. 2008;23:244-250.
18. Wolff D, Herzberg P, Heussner P, et al. Chronic GvHD of the lung significantly impairs quality of life and the activity profile—results of a prospective German multicentre validation trial. *Bone Marrow Transplant*. 2009;43(Suppl 1):P541.
19. Williams KM, Chien JW, Gladwin MT, Pavletic SZ. Bronchiolitis obliterans after allogeneic hematopoietic stem cell transplantation. *JAMA*. 2009;302:306-314.
20. Marras TK, Chan CK, Lipton JH, et al. Long-term pulmonary function abnormalities and survival after allogeneic marrow transplantation. *Bone Marrow Transplant*. 2004;33:509-517.
21. Dudek AZ, Mahaseth H, DeFor TE, Weisdorf DJ. Bronchiolitis obliterans in chronic graft-versus-host disease: analysis of risk factors and treatment outcomes. *Biol Blood Marrow Transplant*. 2003;9:657-666.
22. Chien JW, Duncan S, Williams KM, Pavletic SZ. Bronchiolitis obliterans syndrome after allogeneic hematopoietic stem cell transplantation—an increasingly recognized manifestation of chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2010;16(Suppl 1):S106-S114.
23. Greinix HT, Pohlreich D, Maalouf J, et al. A single-center pilot validation study of a new chronic GVHD skin scoring system. *Biol Blood Marrow Transplant*. 2007;13:715-723.
24. Flowers ME, Apperley JF, van Besien K, et al. A multicenter prospective phase 2 randomized study of extracorporeal photopheresis for treatment of chronic graft-versus-host disease. *Blood*. 2008;112:2667-2674.
25. Jacobsohn DA, Rademaker R, Kaup M, Vogelsang GB. Skin response using NIH consensus criteria vs Hopkins scale in a phase II study for steroid-refractory chronic GVHD. *Bone Marrow Transplant*. 2009;44:813-819.
26. Elad S, Zeevi I, Or R, Resnick IB, Dray L, Shapira MY. Validation of the National Institutes of Health (NIH) scale for oral chronic graft-versus-host disease (cGVHD). *Biol Blood Marrow Transplant*. 2010;16:62-69.
27. Treister NS, Stevenson K, Kim H, Woo SB, Soiffer R, Cutler C. Oral chronic graft-versus-host disease scoring using the NIH consensus criteria. *Biol Blood Marrow Transplant*. 2010;16: 108-114.
28. Perez-Simon JA, Encinas C, Silva F, et al. Prognostic factors of chronic graft- versus-host disease following allogeneic peripheral blood stem cell transplantation: The National Institutes of Health scale plus the type of onset can predict survival rates and the duration of immunosuppressive therapy. *Biol Blood Marrow Transplant*. 2008;14:1163-1171.
29. Shulman HM, Sale GE, Lerner KG, et al. Chronic cutaneous graft-versus-host disease in man. *Am J Pathol*. 1978;91:545-570.
30. Sullivan KM, Shulman HM, Storb R, et al. Chronic graft-versus-host disease in 52 patients: adverse natural course and successful treatment with combination immunosuppression. *Blood*. 1981;57:267-276.
31. Wingard JR, Piantadosi S, Vogelsang GB, et al. Predictors of death from chronic graft-versus-host disease after bone marrow transplantation. *Blood*. 1989;74:1428-1435.
32. Akpek G, Zahurak ML, Piantadosi S, et al. Development of a prognostic model for grading chronic graft-versus-host disease. *Blood*. 2001;97:1219-1226.
33. Akpek G, Lee SJ, Flowers ME, et al. Performance of a new clinical grading system for chronic graft-versus-host disease: a multicenter study. *Blood*. 2003;102:802-809.
34. Arora M, Burns LJ, Davies SM, et al. Chronic graft-versus-host disease: a prospective cohort study. *Biol Blood Marrow Transplant*. 2003;9:38-45.
35. Jagasia MH, Savani BN, Stricklin G, et al. Classic and overlap chronic graft- versus-host disease (cGVHD) is associated with superior outcome after extracorporeal photopheresis (ECP). *Biol Blood Marrow Transplant*. 2009;15:1288-1295.

APPENDIX 1: THE FOLLOWING INDIVIDUALS PARTICIPATED IN THE CONSENSUS PROCESS:

HSCT center(s)	Name of Representatives	Country
Survey and Consensus Conferences:		
Berlin	A. Gerbitz, L. Uharek	Germany
Dresden	A. Kiani	Germany
Freiburg	A. Bertz	Germany
Hamburg	A. Ayuk, U. Bacher, A. Zander	Germany
Hannover	M. Stadler	Germany
Leipzig	B. Basara	Germany
Muenster	M. Stelljes	Germany
Oldenburg	J. Casper	Germany
Regensburg	D. Wolff	Germany
Rostock	S. Hilgendorf	Germany
Ulm	S. v. Harsdorf	Germany
Vienna*	H. Greinix, A. Lawitschka	Austria
Other Participants in Consensus Conference/Discussions:		
Augsburg	C. Schmid	Germany
Berlin	R. Arnold, M. Hildebrandt, J. Kuehl, K. Rieger	Germany
Cologne	C. Scheid	Germany
Duesseldorf	G. Kobbe, R. Meisel	Germany
Dresden	M. Suttorp	Germany
Erlangen	W. Roesler, W. Holter	Germany
Essen	A. Elmaagacli, B. Kremens	Germany
Frankfurt am Main	H. Martin, P. Bader	Germany
Freiburg	B. Strahm	Germany
Giessen	W. Woessmann	Germany
Goettingen	J. Hasenkamp	Germany
Greifswald	G. Doelken	Germany
Hamburg	H. Kabisch	Germany
Hannover	S. Buchholz, K. Sykora	Germany
Heidelberg	T. Luft	Germany
Jena	K. Schilling, K. Kentoche	Germany
Kiel	M. Gramatzki, A. Claviez	Germany
Mainz	R.G. Meyer	Germany
Munich	J. Tischer, M. Albert, I. Luettichau	Germany
Muenster	K. Ehlert	Germany
Nuernberg	S. Dressler, S. Wandt	Germany
Regensburg	R. Andreesen, E. Holler	Germany
Stuttgart	P. Schlegel	Germany
Tuebingen	W. Bethge, P. Schwarze	Germany
Wiesbaden	M. Schleuning, R. Schwerdtfeger	Germany
Wuerzburg	G. Stuhler	Germany
Graz	C. Urban	Austria
Innsbruck†	G. Kropshofer	Austria
Linz	O. Krieger	Austria
Vienna	Z. Kuzmina	Austria
Basel	J. Halter	Switzerland

*In Vienna, 2 separate institutions (St. Anna Childrens' Hospital and Medical University of Vienna) participated in the survey.

†In Innsbruck, only the pediatric HSCT program participated in the consensus process.