National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group Report

Madan Jagasia,[1] Hildegard Greinix,[2] Mukta Arora,[3] Kirsten Williams,[4,5] Daniel Wolff,[6] Edward Cowen,[4] Jeanne Palmer,[7] Dan Weisdorf,[3] Nathaniel Treister,[8] Guang-Shing Cheng,[9] Holly Kerr,[10] Pam Stratton,[11] Rafael Duarte,[12] George McDonald,[9] Yoshihiro Inamoto,[13] Afonso Vigorito,[14] Sally Arai,[15] Manuel Datiles,[16] David Jacobsohn,[5] Theo Heller,[17] Carrie L. Kitko[18], Sandra Mitchell,[4] Paul J Martin,[9] Howard Shulman,[9] Roy Wu,[4] Corey Cutler,[8] Georgia Vogelsang,[16] Stephanie J. Lee,[9] Steven Z. Pavletic,[4] and Mary E.D. Flowers[9]

[1]Vanderbilt University Medical Center, Nashville, TN
[2] Medical University of Vienna, Vienna, Austria
[3] University of Minnesota, Minneapolis, MN
[4] National Cancer Institute, NIH, Bethesda, MD
[5] Children's National Medical Center for Cancer & Blood Disorders, Washington, DC
[6] Department of Internal Medicine III, University of Regensburg, Regensburg, Germany
[7] Hematology Oncology/Blood and Marrow Transplant, Mayo Clinic Arizona, Phoenix, AZ
[8] Dana Farber Cancer Institute, Brigham and Women's Hospital, Boston, MA
[9] Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA
[10] BC Cancer Agency, Vancouver, British Columbia, Canada
[11]National Institute of Child Health and Human Development, NIH, Bethesda, MD
[12] Institut Catala d'Oncologia, Spain
[13] Division of Hematopoietic Stem Cell Transplantation, National Cancer Center Hospital, Tokyo, Japan
[14] Hemocentro Unicamp, Campinas, SP, Brazil
[15] Stanford University Medical Center, Stanford, CA
[16] National Eye Institute, NIH, Bethesda, MD
[17]National Institute of Diabetes and Digestive and Kidney Disease, NIH, Bethesda, MD
[18] Blood and Marrow Transplant Program, University of Michigan Comprehensive Cancer Center, Ann Arbor, MI
[19] Johns Hopkins, Baltimore, MD

Corresponding author: Mary E.D. Flowers, MD, Member, FHCRC; Professor of Medicine, University of Washington; Telephone: (206) 667-5160; Email: mflowers@fhcrc.org

Keywords: Chronic GVHD, NIH, diagnosis, staging

Running title: Diagnosis and staging of chronic Graft-versus-Host Disease

Abstract word count: 147; Manuscript word count: 6,516;

Tables: 2; Figure: 1; Supplemental Figure: 1

References: 51

Abstract

The 2005 National Institute of Health (NIH) Consensus Conference proposed new criteria for diagnosis and scoring the severity of chronic GVHD. The 2014 NIH consensus maintains the framework of the prior consensus with further refinement based on new evidence. Revisions have been made to address areas of controversy or confusion, such as the overlap chronic GVHD subcategory and the distinction between active disease and past tissue damage. Diagnostic criteria for involvement of mouth, eyes, genitalia and lungs have been revised. Categories of chronic GVHD should be defined in ways that indicate prognosis, guide treatment, and define eligibility for clinical trials. Revisions have been made to focus attention on the causes of organ-specific abnormalities. Attribution of organ specific abnormalities to chronic GVHD has been addressed. This paradigm shift provides greater specificity, more accurately measures the global burden of disease attributed to GVHD, and will facilitate biomarker association studies.

Background

Chronic graft-versus-host disease (GVHD) remains a serious and common complication of allogeneic hematopoietic cell transplantation (HCT), occurring in 30% to 70% of patients[1]. The 2-year cumulative incidence of NIH defined chronic GVHD after marrow or peripheral blood HCT from related or unrelated donors is 34% (range, 32% - 35%)[2]. Chronic GVHD is a syndrome of variable clinical features resembling autoimmune and other immunologic disorders such as scleroderma, Sjögren syndrome, primary biliary cirrhosis, wasting syndrome, bronchiolitis obliterans, immune cytopenias, and chronic immunodeficiency[3,4]. The pathophysiology of the chronic GVHD syndrome may involve inflammation, cell mediated immunity, humoral immunity and fibrosis. Clinical manifestations nearly always present during the first year after transplantation, but some cases develop many years after HCT. Manifestations of chronic GVHD may be restricted to a single organ or site or may be widespread, with profound impact on quality of life. Other cases are self-limited and either smolder or resolve without immunosuppressive therapy.

Diagnosing and scoring the severity of chronic GVHD may be challenging for several reasons: limited understanding of the pathophysiology, coexistence of acute GVHD manifestations, previously poorly validated measurement tools/scoring systems and, lack of biomarkers for the diagnosis and assessment of disease activity.

Overall risk profiles for acute GVHD and for chronic GVHD diagnosed per 2005 NIH consensus criteria[5] were similar in a large comparative study[2]. Of interest, risk factors associated with chronic GVHD were not changed after adjustment for prior acute GVHD, suggesting that chronic GVHD is not simply an evolution of preceding acute GVHD[2].

Several retrospective and large prospective studies have validated many aspects of the 2005 NIH Chronic GVHD Diagnosis and Staging Consensus criteria[5] including organ scoring, global severity and GVHD categories[6-21]. Although these criteria represent advancement in the field, many questions remain, including their role in clinical practice, clinical trials biomarker discovery, and regulatory review of new drugs or devices seeking FDA approval. For certain organs and sites, the minimal criteria to diagnose chronic GVHD have not been clearly defined. Other unresolved issues of the 2005 Consensus criteria include confusion about the chronic GVHD subcategories (especially overlap GVHD), the rules for scoring abnormalities (symptoms, signs, diagnostic testing) not due to GVHD and lack of distinction between active disease and a fixed deficit resulting from prior tissue damage[6,22].

The 2014 international NIH Chronic GVHD Diagnosis and Scoring Consensus Working Group that contributed to this document were subdivided into organ specific subgroups. Each subgroup reviewed all new evidence since 2005 and was asked to address controversies and unanswered questions within their specific organ[22]. Their findings were reviewed by all the members of the working group and the steering committee and agreed upon, to establish the 2014 Consensus Criteria.

Purpose of this document

The goals of this consensus document are to revise the 2005 NIH chronic GVHD consensus criteria[5] based on available evidence, to (a) clarify controversies related to the minimal criteria needed to establish the diagnosis in clinical practice; and (b) refine the definition of GVHD sub-categories and organ severity scoring. The changes proposed in this document will help to identify manifestations of the various clinical phenotypes of chronic GVHD at initial diagnosis

and during the subsequent evolution of the disease for the purpose of clinical trials and biomarkers studies needed to advance the field. A summary of the 2014 NIH Chronic GVHD Diagnosis and Scoring Consensus Recommendations is shown below.

Summary of recommendations that are new from the 2005 Consensus[5]

1. Definition of overlap chronic GVHD subcategory has been clarified and specific manifestations of both acute and chronic GVHD have been added to the organ severity scoring form.

2. Diagnostic criteria for organ system involvement have been modified as follows:
   a. Mouth: Hyperkeratotic plaques have been removed as a diagnostic feature.

   b. Eyes: Evaluation by an ophthalmologist is recommended for eye-specific clinical trials. The Schirmer test has been removed from the severity scoring form.

   c. Lungs: Bronchiolitis obliterans syndrome (BOS) diagnostic criteria have been modified to enhance diagnostic sensitivity. BOS that meets the new criteria for lung manifestation and lung biopsy confirming BO are now each defined as diagnostic features.

   d. Genitalia: Signs and symptoms for male have been added and diagnostic criteria for females have been modified.

3. Organ-specific severity scoring has been modified as follows (Figure 1):
    a. Skin: The composite score has been split into two scores to separate the extent of skin involvement (body surface area - BSA) from the specific skin features. Clinical features to be considered in the skin scores have been clarified and rules for the final skin scoring have been added for calculation of global severity.

    b. Mouth: Asymptomatic lichen planus-like features (score 0) has been incorporated.

    c. Eye: Kerato-conjunctivitis sicca (KCS) confirmed by an ophthalmologist in an asymptomatic patient (score 0) has been incorporated. Scoring regarding eye drops criterion is clarified to include only lubricant drops.

    d. Gastrointestinal (GI): Severity of diarrhea has been added to the GI tract severity score.

    e. Liver: Aspartate aminotransferase (AST) is no longer included in liver severity scoring. The cut off values for bilirubin, alanine aminotransferase (ALT) and alkaline phosphatase (ALP) have been revised.

    f. Lungs: The lung function score which included both FEV1 and DLCO has been simplified to include only the FEV1, thus increasing specificity of obstructive lung defects. Rules for final lung scoring have been modified to enhance specificity and for calculation of global severity.

g. Joint: Photographic image-based range of motion (P-ROM)[23] has been added to the joint assessment as an exploratory measure.

h. Genitalia: New criteria are proposed for scoring severity based on signs as an exploratory measure.

i. Other indicators have been removed, including the category of progressive onset, and cardiac manifestations such as conduction defects and coronary artery involvement (Figure 1). Weight loss (not due to gastrointestinal involvement by GVHD) has been added to this section

j. Attributions of abnormalities not due to GVHD have been incorporated in the organ-specific scoring.

4. The evaluator's opinion regarding overall severity of chronic GVHD has been added to the scoring form (Figure 1).

Diagnosis of chronic GVHD

Clinical features determine whether the clinical syndrome of GVHD is considered acute or chronic, not the temporal relationship to transplantation[5]. In the 2005 consensus criteria, the simultaneous presence of acute GVHD features in patients with chronic GVHD was classified as overlap GVHD[5]. Overlap GVHD has been a subject of controversy and confusion (see Differential Diagnosis between Acute and Chronic GVHD in the following section). The

overlap GVHD subcategory has been associated with worse survival compared to the "classic" subcategory (absence of acute GVHD features) of chronic GVHD[9,13,20,24], but not in all studies[7,18]. Hyperbilirubinemia and small intestinal/colonic involvement are known risk factors for increased mortality in chronic GVHD patients (reviewed in [3])[7,25,26]. Based on current knowledge and in light of controversy related to the overlap subcategory including problems identified in clinical practice[22], the 2014 consensus criteria have clarified overlap GVHD and recommends documentation of all clinical features in patients with chronic GVHD that are relevant for prognostication, treatment guidance, response assessment, biomarker studies and clinical trials (see "Differential Diagnosis between Acute and Chronic GVHD" and "Clinical Scoring of Organ Systems" sections below).

Throughout this document, *diagnostic* signs and symptoms refer to those manifestations that establish the presence of chronic GVHD without need for further testing or evidence of other organ involvement. *Distinctive* signs and symptoms of chronic GVHD refer to those manifestations that are not ordinarily found in acute GVHD but are not considered sufficient in isolation to establish an unequivocal diagnosis of chronic GVHD. Additional testing such as a biopsy documenting histological features of chronic GVHD or the presence of distinctive features in another site is needed to establish the diagnosis of chronic GVHD. *Other features or unclassified entities* of chronic GVHD define the rare, controversial, or non-specific features of chronic GVHD that cannot be used to establish the diagnosis of chronic GVHD. *Common* signs and symptoms of chronic GVHD refer to manifestations found in both chronic and acute GVHD (Table 1).

Characteristics of the clinical features that establish the diagnosis of chronic GVHD might not serve as the most appropriate parameters for assessing severity of chronic GVHD. Valid and

reliable diagnostic criteria might not be sufficiently sensitive to change to be useful as treatment-response criteria. Conversely, a sensitive measure of chronic GVHD response might not necessarily serve as an appropriate diagnostic and scoring tool.

The Working Group recommends that the diagnosis of chronic GVHD require at least one diagnostic manifestation of chronic GVHD or at least one distinctive manifestation, with the latter confirmed by pertinent biopsy, laboratory tests, evaluation by a specialist (ophthalmologist, gynecologist) or radiology in the same or other organ, unless stated otherwise. As in acute GVHD, infection and other causes may confound or complicate the differential diagnosis of chronic GVHD and must be excluded (e.g., nail dystrophy due to onychomycosis, herpes simplex or *Candida albicans* infections of the oral cavity, drug toxicity). Diagnostic and distinctive features of chronic GVHD can be found in the skin and appendages, mouth, eyes, genitalia, esophagus, lungs, and connective tissues. Biopsy or other testing is always encouraged and often valuable to confirm the presence of chronic GVHD, but is not always feasible and is not mandatory if the patient has at least one of the diagnostic findings of chronic GVHD (Table 1).

Organ-specific manifestations of chronic GVHD

In all cases, drug reaction, infection, recurrent or new malignancy and other causes must be excluded. Diagnostic clinical or laboratory features sufficient for the diagnosis of chronic GVHD are italicized in the sections below.

Skin

Diagnostic clinical features include *poikiloderma* (e.g., atrophy, pigmentary changes and telangiectasia), *lichen planus-like eruption* (e.g., erythematous/violaceous flat topped papules or plaques with or without surface reticulations or a silvery or shiny appearance), *deep sclerotic features* (e.g., smooth, waxy, indurated skin - "thickened or tight skin", caused by deep and diffuse sclerosis over a wide area generally causing limitation of joint mobility), *morphea-like* superficial sclerotic features (e.g., localized patchy areas of moveable smooth or shiny skin, leathery-like consistency, often with dyspigmentation) or as *lichen sclerosus-like lesions* (e.g., discrete to coalescent gray to white moveable papules or plaques, often with follicular plugs, shiny appearance, and cigarette paper-like wrinkled texture). Severe sclerotic features characterized by thickened, tight, and fragile skin are often associated with poor wound-healing, inadequate lymphatic drainage, and skin ulcers from minor trauma.

Depigmentation and papulosquamous lesion are "distinctive" features of chronic GVHD (i.e., not seen in acute GVHD, but not sufficiently unique to be considered diagnostic of chronic GVHD). These features contribute to the diagnosis of chronic GVHD in combination with biopsy or laboratory confirmation of GVHD in skin or another organ. Sweat impairment and intolerance to temperature change from loss of sweat glands are seen in chronic GVHD, and are considered in the "other" feature category along with other manifestations such as ichthyosis, keratosis pilaris, hypopigmentation and hyperpigmentation (Table 1). These "other" features cannot be used to establish the initial diagnosis of chronic GVHD. Skin manifestations found in both acute and chronic GVHD include erythema, maculopapular rash and pruritus are categorized as "common" features. The presence of one or more of the "common" features

(without a diagnostic criterion in another organ) cannot be used to establish the initial diagnosis of chronic GVHD.

Assessment of extent and severity of skin chronic GVHD is complex because some clinical features may reflect past 'damage' (hypo- and hyper- pigmentary changes) or sequelae of long-standing fibrosis (i.e., fixed joint contractures after several years of deep sclerosis). Assessment of disease activity is difficult in patients with poikiloderma (atrophic skin, hyperpigmentation, hypopigmentation and telangiectasia) when smoldering ill-defined erythema is admixed with pigmentary changes. Pigmentary change alone (seen in poikiloderma, or more commonly as simple post-inflammatory pigmentary change and not representing active GVHD) is *not* included in the percentage of BSA skin score calculation (See Table 1/Figure 1). Erythema, a "common" feature, is included in the BSA skin score calculation as it generally represents inflammation associated with active GVHD. The erythema only component of poikiloderma manifestation is considered in the BSA skin score calculation but it may be difficult to quantify since it is admixed with pigmentary changes.

Nails

Dystrophy consisting of longitudinal ridging, nail splitting or brittleness, onycholysis, pterygium unguis, and nail loss (usually symmetric and affecting most nails) are distinctive signs of chronic GVHD.

Hair

Distinctive features of chronic GVHD include new scarring or non-scarring scalp alopecia (not due to chemotherapy or radiotherapy) and loss of body hair. Other characteristics seen with chronic GVHD include premature graying, thinning, or brittleness.

Mouth

Diagnostic features of oral chronic GVHD include *lichen planus-like changes*, characterized by hyperkeratotic white lines and lacy-appearing lesions and plaque-like changes affecting the oral mucosa. Changes are typically observed in the buccal mucosa and tongue, although all intraoral surfaces and the vermillion lip may be involved. These diagnostic white changes may be observed with or without associated erythema or ulcerations, which are not considered "diagnostic" features. The presence of isolated hyperkeratotic plaques without lichen planus-like changes, so called leukoplakia is no longer considered a diagnostic criterion, since these lesions should be considered a separate clinical entity that may imply malignant potential. Decreased range of motion of the jaw secondary to skin sclerosis should be assessed according to skin criteria, and is no longer considered as diagnostic criterion in the oral section. Distinctive features of chronic GVHD include xerostomia (dryness), mucoceles, mucosal atrophy, ulcers and pseudomembranes, but infectious pathogens such as yeast or herpes virus, and secondary malignancy must be excluded. Manifestations common to both acute and chronic GVHD include gingivitis, mucositis, erythema and pain. Figure 1 details the scoring and incorporates asymptomatic oral chronic GVHD as a diagnostic feature.

Eyes

Distinctive manifestations of chronic GVHD include new onset of dry, "gritty", or painful eyes, cicatricial conjunctivitis, keratoconjunctivitis sicca (KCS) and confluent areas of punctate keratopathy. Other features include photophobia, periorbital hyperpigmentation, and blepharitis (erythema of the eye lids with edema, telangiectasia of lid margin). New ocular sicca documented by low Schirmer's test with a mean value of $\leq$ 5 mm at 5 minutes (preferably with confirmation of normal value at an established baseline) or a new onset of keratoconjunctivitis sicca by slit lamp exam with mean Schirmer test values of 6 to 10 mm (preferably with confirmation of normal values at an established baseline) not due to other causes is sufficient for the diagnosis of ocular chronic GVHD for purpose of treatment and for clinical trials designed specifically for KCS. Patients with ocular symptoms prior to transplant should be evaluated by an ophthalmologist for assessment of ocular surface including presence of KCS, conjunctival scaring and inflammation. Baseline evaluation post-transplant (approximately day 100) is strongly encouraged by some experts[27,28]. Figure 1 details the scoring and incorporates asymptomatic ocular chronic GVHD. The scoring of ocular involvement includes the number of times an individual has to use lubricant eye drops each day. The international consensus guidelines on ocular GVHD, have proposed a more detailed scoring schema which involves comprehensive ophthalmological evaluation including pre-transplant evaluation[28]. These remain to be validated and should be considered in clinical trials addressing ocular involvement. Schirmer's test may be useful for diagnosis of ocular GVHD, but the numerical value is not useful for follow-up of ocular GVHD due to poor correlation with symptom change[15]. For this reason, Schirmer's test value has been removed from the scoring form in the current recommendation (Figure 1).

Genitalia

Chronic GVHD of the genital tract (female and male) is often associated with oral chronic GVHD[29,30]. Diagnostic features of genitalia chronic GVHD include *lichen planus-like features, lichen sclerosus-like features, vaginal scarring, clitoral/labial agglutination (females)*, *phimosis and urethral/meatus scarring or stenosis (males)***.**

Genital examination is recommended, even in asymptomatic patients (female and male), especially if signs of chronic GVHD are present in the mouth. If a gynecologist is unavailable, external examination may be performed, but, in this instance, vaginal scarring may be missed (Supplemental Figure 1).

Female genitalia: The vulva and vagina may be affected by chronic GVHD. Symptoms may include dryness, burning, pruritus, pain to touch, dysuria and dyspareunia either with penile insertion or deep penetration leading to sexual dysfunction. Signs of genital chronic GVHD may include patchy or generalized erythema, tenderness on cotton tipped applicator palpation of vestibular gland openings or vulvar mucosa, mucosal erosions or fissures, lace-like leukokeratosis, labial resorption, labial fusion or clitoral hood agglutination, fibrinous vaginal adhesions, circumferential fibrous vaginal banding, vaginal shortening, synechiae, dense sclerotic changes, and complete vaginal stenosis[29,31-34].

Male genitalia: Manifestations of chronic GVHD may be under recognized and underreported in men. The glans penis and the urethra or meatus may be affected. Patients may report painful sexual intercourse, and a burning sensation. Genital signs of GVHD include non-infectious balanoposthitis, lichen sclerosis-like or lichen planus-like features, phimosis or urethra or meatus scaring or stenosis[35,36].

Gastrointestinal tract (GI)

Diagnostic features include *esophageal web*, *stricture, or concentric rings* documented by endoscopy, or barium contrast radiograph. Chronic GVHD may be associated with pancreatic atrophy and exocrine insufficiency leading[37] to malabsorption which often improves with oral pancreatic enzyme supplementation. Manifestations common to both acute and chronic GVHD include anorexia, nausea, vomiting, diarrhea, weight loss, and failure to thrive. These symptoms can be due to non-GVHD causes such as drug side effect, motility disorders or infections. Wasting syndrome may be a manifestation of chronic GVHD but is often multifactorial (e.g., decreased caloric intake, poor intestinal absorption of macronutrients, increased resting energy expenditures and hypercatabolism). Unintentional weight loss occurring over a three month period should be documented irrespective of causality in clinical trials, unless definitive causality other than GVHD is identified. Endoscopic findings of gastrointestinal mucosal edema and erythema or focal erosions with histologic changes of apoptotic epithelial cells and crypt cell dropout are manifestations of acute but not chronic GVHD.

Liver

There are no hepatic manifestations that are either distinctive or diagnostic of chronic GVHD. Liver GVHD can also be accompanied by clinical manifestations of acute GVHD, with or without manifestations of chronic GVHD. Other potential causes of liver disease occurring more than day 100 after HCT, include viral infections, biliary obstruction, drug toxicity and other less common disorders (i.e., nonalcoholic steatohepatitis, and others). GVHD after day 100 can present in two ways. One resembles acute hepatitis (steeply rising serum ALT, with or

without jaundice), almost always after tapering of immunosuppressive drugs or after donor lymphocyte infusion. This presentation requires a prompt diagnosis and treatment intervention, and may need a liver biopsy in the absence of chronic GVHD in other organ. The other resembles a slowly progressive cholestatic disorder with elevation of serum alkaline phosphatase and gamma-glutamyl transpeptidase, followed by jaundice. Acute hepatitis and progressive cholestatic features are included in the "common" category. The liver has no clinical features in the "other" category.

Lungs

Historically, the only diagnostic pulmonary manifestation of chronic GVHD was biopsy-proven bronchiolitis obliterans (*BO)*. However, because biopsy is invasive and associated with risk of bleeding and other complications, experts now endorse the diagnosis of bronchiolitis obliterans syndrome (BOS) using pulmonary functions testing (PFT)[38,39]. BOS is characterized by the new onset of an obstructive lung defect. Clinical manifestations may include dyspnea on exertion, cough or wheezing; however many patients are often asymptomatic early in the disease process. For this reason, screening PFTs is recommended at day 100 posttransplant, at initial diagnosis of chronic GVHD, at one year after transplant and at least, at 6-month intervals for the first two years after the initial diagnosis of chronic GVHD. More often PFTs monitoring is recommended in patients diagnosed with BOS and in those with significant decline in lung volumes but not yet meeting the criteria for BOS (see upcoming supportive care and ancillary care NIH consensus document). Pneumothorax, pneumomediastinum or subcutaneous emphysema are rare and often associated with advanced disease. Restrictive pulmonary function abnormalities are not characteristics of BOS but may reflect extra-pulmonary restriction (leading to false reduction of FEV1), secondary to advanced sclerotic GVHD of the chest wall or other

intrapulmonary processes not related to GVHD, such as cryptogenic organizing pneumonia or pulmonary fibrosis. Further investigation beyond simple pulmonary testing is needed to evaluate these complex problems.

In patients with previously established diagnosis of chronic GVHD, *bronchiolitis obliterans syndrome* (BOS) is a diagnostic feature of lung chronic GVHD when all of the following criteria are met:

(1) FEV1/VC < 0.7 or the 5th percentile of predicted.

   a. FEV1= Forced Expiratory Volume in 1 second.

   b. VC= Vital Capacity (Forced Vital Capacity "FVC" or Slow Vital Capacity "SVC", whichever is greater).

   c. The 5th percentile of predicted is equivalent to the lower value of predicted confidence interval.

   d. For pediatric patients or elderly populations, use < predicted confidence interval using NHANESIII calculations[40]

(2) %FEV1 < 75% of predicted with ≥ 10% decline over less than 2 years. **%**FEV1 should not correct to > 75% with albuterol and the rate of decline for the corrected values should still remain at ≥ 10% decline over 2 years.

(3) Absence of infection in the respiratory tract, documented with investigations directed by clinical symptoms, such as radiologic studies (radiographs or computed tomographic scans) or microbiologic cultures (sinus aspiration, upper respiratory tract viral screen, sputum culture, bronchoalveolar lavage).

(4) One of the two supporting features of BOS:

a. Evidence of air trapping by expiratory CT or small airway thickening or bronchiectasis by high-resolution chest computed tomography OR

b. Evidence of air trapping by PFTs: RV > 120% (Residual Volume) or RV/TLC elevated RV/TLC outside of the confidence interval (RV/Total Lung Capacity).

If BOS presents as the only clinical manifestation in a patient without a prior established diagnosis of chronic GVHD, a distinctive manifestation must be present and a lung biopsy should be considered to confirm the diagnosis of chronic GVHD.

The current recommended work-up for BOS includes PFT testing and expiratory CT. Because a new diagnostic technique for BOS termed parametric response mapping is currently under investigation, a high resolution (helical) CT of inspiration and expiration is encouraged if available. This technique will permit visual representation of lung affected by obstructive disease (BOS) versus lung tissue with normal aeration or restrictive disease and may become a valuable measure in the future[41].

Other entities that are currently not diagnostic or distinctive of lung chronic GVHD, but remain areas of active investigation include: (1) cryptogenic organizing pneumonia (COP) (formerly known as bronchiolitis obliterans organizing pneumonia), and (2) progressive restrictive lung disease (in the absence of extra pulmonary causes). These unclassified entities have been placed in the "other" category in Table 1. There are no "common" pulmonary features of GVHD.

Musculoskeletal system

Diagnostic features include *fascial involvement* often affecting the forearms or legs and often associated with sclerosis of the overlying skin and subcutaneous tissue. Fascial involvement may develop without overlying sclerotic changes of the skin, and can result in *joint stiffness or contractures* when present near joints. Early fasciitis may present with pain and swelling with or without erythema. *Fasciitis* is detected on examination by stiffness, restricted range of motion (e.g., often decreased dorsal wrist flexion or inability to assume a Buddha prayer posture), edema of extremities with or without erythema (early sign), peau d'orange (edematous skin with prominent pores resembling the surface of an orange) or *joint contractures* (late complications). Clinical myositis with muscles tenderness and elevated muscle enzymes in the blood is a distinctive but non-diagnostic manifestation of chronic GVHD. Myositis may present as proximal myopathy, but this complication is rare and does not explain the frequent complaints of severe cramps. Evaluation of myositis includes electromyography and measurement of creatinine phosphokinase or aldolase. Muscle/sural nerve biopsies should be considered in the absence of other manifestations of GVHD and to rule out other causes of myositis. Arthralgia and "true" arthritis are uncommon and are occasionally associated with the presence of autoantibodies.

Hematopoietic and immune systems

Hematopoietic and immunological abnormalities are frequently associated with chronic GVHD but cannot be used to establish the diagnosis of chronic GVHD. Cytopenias may result from stromal damage or autoimmune processes. Lymphopenia ($\leq$500/μl), eosinophilia ($\geq$ 500/ μl), hypogammaglobulinemia or hypergammaglobulinemia may be present. Autoantibodies may develop with autoimmune hemolytic anemia and idiopathic thrombocytopenic purpura.

Thrombocytopenia (<100,000/μl) at the time of chronic GVHD diagnosis has been associated with a poor prognosis.

Other findings

Serositis (pericardial or pleural effusions or ascites), peripheral neuropathy, myasthenia gravis, nephrotic syndrome, membranous glomerulonephritis, Raynaud phenomenon and cardiac involvement have been attributed to chronic GVHD, but these manifestations are rare. For these entities, attribution to chronic GVHD is often a diagnosis of exclusion.

Differential Diagnosis between Acute and Chronic GVHD

As in the 2005 consensus criteria, the 2014 consensus recognizes two main categories of GVHD (acute and chronic). The broad category of acute GVHD includes (1) classic acute GVHD (erythema, maculopapular rash, nausea, vomiting, anorexia, profuse diarrhea, ileus or cholestatic liver disease) occurring within 100 days after transplantation or DLI (without diagnostic or distinctive signs of chronic GVHD) and (2) persistent, recurrent or late acute GVHD: features of classic acute GVHD without diagnostic or distinctive manifestations of chronic GVHD occurring beyond 100 days of transplantation or DLI (often seen during the taper or after withdrawal of immune suppression).

In the 2005 criteria, the broad category of chronic GVHD included (1) classic chronic GVHD without features characteristic of acute GVHD and (2) and an overlap syndrome where features of chronic and acute GVHD appear together. Clarification of the definition of "overlap"

subcategory of chronic GVHD is now provided to address problems identified in clinical practice in applying this terminology[22]. The term "overlap" refers to the presence of acute GVHD manifestation in a patient with a diagnosis of chronic GVHD. Manifestations of acute GVHD can be present at initial diagnosis of chronic GVHD or can develop after the diagnosis of chronic GVHD and may recur with or without resolution of prior chronic GVHD manifestations. Overlap subcategory can be transient, often depends on the degree of immunosuppression, and is subject to changes during the disease course. Many patients who present with "overlap" chronic GVHD resolve the acute features while the chronic GVHD features persist. Similarly patients with classic chronic GVHD may develop acute GVHD features when immunosuppression is tapered. The 2014 chronic GVHD consensus recommends documentation of all specific manifestations (acute and chronic) when scoring organ severity at initial and at any time after the diagnosis of chronic GVHD (Figure 1). Advantages of capturing specific manifestations over simply referring to "overlap" include: better description of the chronic GVHD syndrome at any time, useful to guide treatment, identify manifestations of prognostic value, helpful in identify candidates for specific clinical trials and for biomarkers and biology studies. Specific manifestations are shown in Figure 1 and are discussed in the section about scoring below. For example, skin sclerosis and fasciitis manifestations have been separated from BSA calculations that are more applicable to other manifestations such as erythema. Severity of diarrhea has been added to the GI tract scoring. Liver scoring was modified to reflect the biochemical liver abnormalities that are affected in early versus later (or more severe) phases of GVHD.

In the absence of features fulfilling the definition of chronic GVHD, the persistence, recurrence or new onset of characteristic skin, gastrointestinal tract or liver abnormalities should be classified as acute GVHD regardless of the time after transplantation. With appropriate

stratification, however, patients with persistent, recurrent or late acute GVHD may be included in clinical trials together with patients who have NIH chronic GVHD[2].

Clinical Scoring of Organ Systems

Modifications have been made to the 2005 consensus organ scoring system based on available evidence, or lack thereof, and to address concerns raised by investigators and in clinical practice[22]. Figure 1 shows the consensus scoring system for individual organs. Several considerations explain the selection of the features for the proposed scoring system versus the response criteria discussed in a separate article. (1) Scoring criteria are intended for baseline or cross-sectional use, while response criteria are intended for longitudinal evaluation in therapeutic trials. (2) In general, scoring measures have been designed so that they can be easily performed by general practitioners (non-transplant physician and nurses). Two organ systems, eyes and female genitalia (Supplemental Figure 1) are best assessed by an organ-specific consultant. By design, the only required laboratory testing needed to complete the scoring table is measurement of liver tests. Lung scoring is preferentially determined by pulmonary function test, when available, but symptoms may be substituted if PFTs is not available. (3) The broad scoring categories help to classify patients and provide immediate, clinically meaningful information summarizing the disease extent and severity. (4) The scoring system does not attempt to distinguish between disease activity (inflammation and apoptosis or target cells) and fixed anatomic deficits from past tissue injury, but now incorporates the attribution of abnormalities not due to chronic GVHD. (5) In organ systems, with two possible scores (e.g. skin) the higher

score is used for calculating global severity. FEV1 obtained from pulmonary function test (PFT) supersedes the clinical scoring in lung. (6) Sites or organ with unequivocal documentation of attribution other than GVHD cannot be evaluated and are not scored in computing the overall severity, but the data are incorporated in the scoring form (Figure 1). For example, 12.5% BSA skin rash entirely due to varicella zoster is scored as 0 for skin, shortness of breath after walking on flat ground due to lobar pneumonia is scored 0 for lung, FEV1 of 60% is scored 0 if is unchanged from the pre-transplant FEV1 value. We anticipate that patients will often have multifactorial etiologies to explain the abnormality present (e.g. shortness of breath in a patient with established BOS and now with worsening FEV1 due to superimposed viral bronchiolitis). In these instances, the abnormality is scored as if the entire deficit is due to GVHD. This inherent limitation of the scoring system is unavoidable, until better quantitative tests to ascertain abnormality solely due to chronic GVHD are available.

Organ sites considered for scoring include skin, mouth, eyes, gastrointestinal tract, liver, lungs, joints and fascia, and the genital tract. Each organ or site is scored according to a 4-point scale (0-3) with 0 representing no involvement and 3 reflecting severe impairment. In addition, performance status is captured on a 0 to 3 scale, and check boxes note the presence or absence of other specific manifestations.

The current consensus document proposes changes to the 2005 consensus scoring system for some organs as follows (Figure 1):

1. Skin: The composite score is now split into two scores to document the extent of skin involvement (BSA) and the specific skin features separately. Clinical features to be considered in the skin scores have been clarified. The higher of the two scores is to be used for computation of the global severity.

2. Mouth: Lichen planus-like features in asymptomatic patients (score 0) are now incorporated.
3. Eye: Keratoconjunctivitis sicca (KCS) confirmed by an ophthalmologist in an asymptomatic patient (score 0) is now incorporated. Scoring regarding the requirement of eye drops is clarified to include only lubricant drops. Schirmer's test value has been removed from the scoring form.
4. Gastrointestinal (GI): The severity of diarrhea is now incorporated as an additional feature in the GI tract severity scoring system. Weight loss due to gastrointestinal GVHD is captured under the GI tract.
5. Genitalia: Scoring is now based on severity of the signs instead of symptoms, based on limited available data[29,31,35,36] and opinion of experts (supplemental Figure 1 represents an exploratory measure to be completed by specialist or trained practitioners). Female genital GVHD is not scored if a practitioner is unable to examine the patient.
6. Liver: Scoring is based on increments in values for total serum bilirubin, alanine aminotransferase (ALT), and alkaline phosphatase (ALP). Aspartate aminotransferase (AST) is no longer considered for the scoring.
7. Lungs: Lung function score, which used both FEV1 and DLCO, was simplified to FEV1 values alone, thus improving specificity. The rule for the final lung scoring has been changed such that the FEV1 score should be used in cases with discrepancy between symptoms and FEV1 scores.

8. Joint: Photographic-range of motion (P-ROM)[23] has been added to joint assessment as an exploratory measure and should not be included in the calculation of global severity (Figure 1).
9. Other indicators, clinical manifestations or complications related to chronic GVHD have been simplified. These include the removal of progressive onset, cardiomyopathy, cardiac conduction defects and coronary artery involvement. Weight loss (measured over previous 3 months) due to causes other than GI tract GVHD is now captured under this section (Figure 1).

The form shown in Figure 1 should be completed based on an assessment of current status without consideration of past manifestations or the causes for the abnormality in each organ. Abnormalities with unequivocal causes other GVHD are annotated in scoring each organ or site. This change will help to address some of the controversies and confusion raised by investigators[22]. Furthermore, identification of abnormalities not due to GVHD will help in the selection of patients for biomarker studies of chronic GVHD and clinical trials. We realize that abnormality may have a multifactorial etiology. In those instances, the organ should be scored as if the entire abnormality is due to GVHD.

Global Scoring of Chronic GVHD

The fundamentals of the global scoring of chronic GVHD remain unchanged from 2005 NIH consensus criteria[5]. Several studies have shown that the 2005 NIH global severity score baseline predicts overall survival and non-relapse mortality[11,18,42] and some elements of the score have been validated with patient reported quality of life measures[10,43].

Eight organ sites (skin, eyes, gastrointestinal tract, liver, lungs, joint and fasciae, and genital tract) are considered for calculating global score. Elements included in the proposed global scoring include both the number of organs or sites involved and the severity score within each affected organ. Performance status scoring is not incorporated into the global scoring system. The global descriptions of mild, moderate, and severe were chosen to reflect the degree of organ impact and functional impairment due to chronic GVHD. Although scoring is often used at the time of initial diagnosis, evaluating the clinical score periodically during the course of chronic GVHD may revise prognostic expectations and better describe the current severity of chronic GVHD. It is important to note that change in global scoring system over time is not synonymous with response. The global scoring system can be applied only after the diagnosis of chronic GVHD is confirmed by either (1) presence of a diagnostic feature or, if a diagnostic feature is not present, (2) at least one distinctive manifestation of chronic GVHD with the diagnosis supported by histologic, radiologic or laboratory evidence of GVHD from any site or by a distinctive clinical manifestation in another site. Table 2 outlines the computation of the chronic GVHD global severity scoring which is categorized as mild, moderate or severe.

The current consensus incorporates asymptomatic organ manifestation (e.g. asymptomatic oral chronic GVHD). These do not affect the global scoring of chronic GVHD, since the recorded score is still 0. Attribution of abnormalities to causes other than chronic GVHD could have an impact in the global scoring. For instance, if a patient has a score of ≥ 1 in an organ and if the abnormality is explained entirely and unequivocally by a non-GVHD cause, the organ is scored as zero for calculation of the global severity only. The capture of the potential confounders in the organ scoring (attribution due to other causes than chronic GVHD) will correct any overestimation of organ involvement[11,42] and improve the specificity of the scoring

system. These changes are supported by the results of a recent prospective study evaluating the impact of cofounders in the organ scoring and in the global severity of chronic GVHD, and showed that approximately 40% of abnormalities in at least one organ were unequivocally explained to causes other than chronic GVHD resulting in a modest downgrade of global severity after the confounder was taken into account[44]. As outlined previously, if the abnormality in an organ is multifactorial, the organ is scored as if the entire deficit is due to GVHD.

Indications for systemic therapy

Symptomatic mild chronic GVHD may often be managed with local therapies alone (e.g. topical corticosteroids for the skin involvement). In patients with chronic GVHD that involves three or more organs or with a score of 2 or greater in any single organ, however, systemic immunosuppressive therapy should be considered. In some organ sites (mouth, eyes, genital tract), aggressive local therapy alone may be reasonable, as response to systemic therapy may be suboptimal or may not warrant the risk. Co-morbidities and infections may also modify decisions regarding the time and intensity of therapy. Good medical practice and judgment dictate flexibility in this recommendation. Comprehensive monitoring for early detection of insidious disease progression in other sites is mandatory when management relies entirely on local therapy. Early intervention with effective systemic therapy can prevent progression to severe chronic GVHD. Effective immune modulating therapy can ameliorate clinical manifestations and prolong survival. In patients with newly diagnosed chronic GVHD who are already receiving immune suppressive medications, the dosage may be increased or other agents can be added. Chronic GVHD itself and systemic immunosuppressive therapy, both impair immune defenses.

Therefore patients should receive infection-prevention measures as outlined in the forthcoming Ancillary Therapy and Supportive Care working group document.

Assessment of risk of transplant related mortality (TRM)

Chronic GVHD is one of the major causes of late TRM after allogeneic HCT. Prospective studies using the 2005 criteria have shown that the skin score, lung score and gastrointestinal score each predict the risk of TRM[8,10,16,42]. Previous studies have identified several factors associated with an increased risk of TRM among patients with chronic GVHD including, involvement of multiple organs or sites, decreased clinical performance score, thrombocytopenia (platelet count <100,000/μL) at the time of diagnosis, progressive onset of chronic GVHD from prior acute GVHD (or steroid dose at onset of chronic GVHD), hyperbilirubinemia and a higher percentage of skin involvement at the time of diagnosis, and others[2,14,25,45-51]. The characteristics consistently associated with an increased risk of late TRM among patients with chronic GVHD are thrombocytopenia and progressive onset of chronic GVHD from acute GVHD.

The consensus guidelines for assessment of chronic GVHD severity summarized in this document can be used in making decisions about treatment and enrollment in clinical trials. The goals of treatment for chronic GVHD are to relieve symptoms, control disease activity and prevent damage and disability. As a general rule, the intensity of treatment should be calibrated to the extent and severity of disease manifestations. Patients with mild or asymptomatic manifestations limited to a single organ or site can often be managed with close observation or topical treatment, or by slowing the taper of prophylactic immunosuppressive treatment. Those with more severe manifestations or involvement of multiple organs or sites typically require

systemic treatment. Although it is commonly assumed that systemic treatment might improve survival, previous randomized trials have not demonstrated such a benefit, and some studies have shown statistically significant differences or trends indicating worse survival with intensive immunosuppressive treatment. Therefore, chronic GVHD should be managed with the lowest amount of treatment needed to control the disease until immunological tolerance eventually emerges. Therapeutic interventions that facilitate tolerance induction remains an unmet clinical need.

ACKNOWLEDGMENTS

This project was supported by the National Institutes of Health's (NIH's) National Cancer Institute, Center for Cancer Research, Intramural Research Program and Division of Cancer Treatment and Diagnosis, Cancer Therapy Evaluation Program; Office of Rare Disease Research, National Center for Advancing Translational Sciences; Division of Allergy, Immunology and Transplantation, National Institute of Allergy and Infectious Disease; National Heart Lung and Blood Institute, Division of Blood Diseases and Resources. The authors want to acknowledge the following individuals and organizations that, by their participation, made this project possible: American Society for Blood and Marrow Transplantation, Center for International Bone and Marrow Transplant Research, US Chronic GVHD Consortium (supported by ORD and NCI), German-Austrian-Swiss GVHD Consortium, National Marrow Donor Program, the Health Resources and Services Administration, Division of Transplantation, US Department of Human Health and Services, Canadian Blood and Marrow Transplant Group, European Group for Blood and Marrow Transplantation, Pediatric Blood and Marrow Transplant Consortium, and the representatives of the Brazilian Chronic GVHD consortium (Drs. Maria

Claudia Moreira, Marcia and Vaneuza Funke) and Deutsche José Carreras Leukämie-Stiftung. The authors would like to thank Dr. Joseph Antin (Dana Farber Cancer Center, Boston, MA, USA) and Dr. Gerard Socie (University Paris VII & AP-HP, Hospital Saint Louis, Paris, France) for their critical review of the manuscript. The organizers are in debt to patients and patient and research advocacy groups, who made this process much more meaningful by their engagement. Acknowledgement goes to the Meredith Cowden GVHD foundation for facilitating the initial planning meeting in Cleveland in November of 2013 in conjunction with the National GVHD Symposium. The project group also recognizes the contributions of numerous colleagues in the field of blood and marrow transplantation in the US and internationally, medical specialists and consultants, the pharmaceutical industry, and the NIH and US Food and Drug Administration professional staff for their intellectual input, dedication, and enthusiasm on the road to completion of these documents. For their expert contributions to this 2014 NIH Consensus Diagnosis and Staging Working Group document special acknowledgements go to Drs. Mark Schubert, Fred Hutchinson Cancer Research Center in Seattle, Tina Dietrich-Ntoukas, University of Regensburg, Janine Clayton, National Eye Institute, NIH, Bethesda, and Robert Knobler, University of Vienna.

APPENDIX: NATIONAL INSTITUTES OF HEALTH CONSENSUS-DEVELOPMENT PROJECT ON CRITERIA FOR CLINICAL TRIALS IN CHRONIC GVHD STEERING COMMITTEE

Members of this committee included: Steven Pavletic, Georgia Vogelsang and Stephanie Lee (project chairs), Mary Flowers and Madan Jagasia (Diagnosis and Staging), David Kleiner and Howard Shulman (Histopathology), Kirk Schultz and Sophie Paczesny (Biomarkers), Dan

Couriel and Paul Carpenter (Ancillary and Supportive Care), Paul Martin and Corey Cutler (Design of Clinical Trials), Kenneth Cooke and David Miklos (Chronic GVHD Biology subcommittee), Roy Wu, William Merritt, Linda Griffith, Nancy DiFronzo, Myra Jacobs, Susan Stewart, Meredith Cowden (members).

References

1. Lee SJ, Flowers MED. Recognizing and managing chronic graft-versus-host disease. In: Gewirtz, AM, Muchmore, EA, Burns, LJ (eds.). Hematology 2008: American Society of Hematology Education Program Book. (American Society of Hematology: Washington, DC, 2008, pp. 134-141.

2. Flowers MED, Inamoto Y, Carpenter PA, Lee SJ, Kiem H-P, Petersdorf EW *et al.* Comparative analysis of risk factors for acute graft-versus-host disease and for chronic graft-versus-host disease according to National Institutes of Health consensus criteria. Blood 2011; 117: 3214-3219.

3. Lee SJ, Vogelsang G, Flowers MED. Chronic graft-versus-host disease. Biol Blood Marrow Transplant 2003; 9: 215-233.

4. Flowers MED, Vogelsang GB. Clinical manifestations and natural history. In: Vogelsang, GB, Pavletic, SZ (eds.). Chronic Graft Versus Host Disease: Interdisciplinary Management. (Cambridge University Press: New York, NY, 2009, pp. 56-69.

5. Filipovich AH, Weisdorf D, Pavletic S, Socie G, Wingard JR, Lee SJ *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: 945-956.

6. Baird K, Steinberg SM, Grkovic L, Pulanic D, Cowen EW, Mitchell SA *et al.* National Institutes of Health chronic graft-versus-host disease staging in severely affected patients: organ and global scoring correlate with established indicators of disease severity and prognosis. Biol. Blood. Marrow. Transplant. 2013; 19: 632-639.

7. Vigorito AC, Campregher PV, Storer BE, Carpenter PA, Moravec CK, Kiem H-P *et al.* Evaluation of NIH consensus criteria for classification of late acute and chronic GVHD. Blood 2009; 114: 702-708.

8. Jacobsohn DA, Kurland BF, Pidala J, Inamoto Y, Chai X, Palmer JM *et al.* Correlation between NIH composite skin score, patient reported skin score, and outcome: results from the Chronic GVHD Consortium. Blood 2012; 120: 2545-2552.

9. Pidala J, Vogelsang G, Martin P, Chai X, Storer B, Pavletic S *et al.* Overlap subtype of chronic graft-versus-host disease is associated with an adverse prognosis, functional impairment, and inferior patient-reported outcomes: a Chronic Graft-versus-Host Disease Consortium study. Haematologica 2012; 97: 451-458.

10. Pidala J, Kurland B, Chai X, Majhail N, Weisdorf DJ, Pavletic S *et al.* Patient-reported quality of life is associated with severity of chronic graft-versus-host disease as measured by NIH criteria: report on baseline data from the Chronic GVHD Consortium. Blood 2011; 117: 4651-4657.

11. Arai S, Jagasia M, Storer B, Chai X, Pidala J, Cutler C *et al.* Global and organ-specific chronic graft-versus-host disease severity according to the 2005 NIH Consensus Criteria. Blood 2011; 118: 4242-4249.

12. Pidala J, Kurland BF, Chai X, Vogelsang G, Weisdorf DJ, Pavletic S *et al.* Sensitivity of changes in chronic graft-versus-host disease activity to changes in patient-reported quality of life: results from the Chronic Graft-versus-Host Disease Consortium. Haematologica 2011; 96: 1528-1535.

13. Arora M, Pidala J, Cutler CS, Chai X, Kurland B, Jacobsohn DA *et al.* Impact of prior acute GVHD on chronic GVHD outcomes: a chronic graft versus host disease consortium study. Leukemia 2013; 27: 1196-1201.

14. Kuzmina Z, Eder S, Bohm A, Pernicka E, Vormittag L, Kalhs P *et al.* Significantly worse survival of patients with NIH-defined chronic graft-versus-host disease and thrombocytopenia or progressive onset type: results of a prospective study. Leukemia 2012; 26: 746-756.

15. Inamoto Y, Chai X, Kurland BF, Cutler C, Flowers ME, Palmer JM *et al.* Validation of measurement scales in ocular graft-versus-host disease. Ophthalmology 2012; 119: 487-493.

16. Pidala J, Chai X, Kurland BF, Inamoto Y, Flowers ME, Palmer J *et al.* Analysis of gastrointestinal and hepatic chronic graft-versus-host disease manifestations on major outcomes: a Chronic Graft-Versus-Host Disease Consortium study. Biol. Blood. Marrow. Transplant. 2013; 19: 784-791.

17. Jagasia MH, Abonour R, Long GD, Bolwell BJ, Laport GG, Shore TB *et al.* Palifermin for the reduction of acute GVHD: a randomized, double-blind, placebo-controlled trial. Bone Marrow Transplant. 2012; 47: 1350-1355.

18. Cho B-S, Min C-K, Eom K-S, Kim Y-J, Kim H-J, Lee S *et al.* Feasibility of NIH consensus criteria for chronic graft-versus-host disease. Leukemia 2009; 23: 78-84.

19. Arora M, Nagaraj S, Witte J, Defor TE, MacMillan M, Burns LJ *et al.* New classification of chronic GVHD: added clarity from the consensus diagnoses. Bone Marrow Transplant. 2009; 43: 149-153.

20. Pérez-Simón JA, Encinas C, Silva F, Arcos MJ, Díez-Campelo M, Sánchez-Guijo FM *et al.* Prognostic factors of chronic graft-versus-host disease following allogeneic peripheral blood stem cell transplantation: the National Institutes 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.

21. Herzberg PY, Heussner P, Mumm FHA, Horak M, Hilgendorf I, von Harsdorf S *et al.* Validation of the human activity profile questionnaire in patients after allogeneic hematopoietic stem cell transplantation. Biol Blood Marrow Transplant 2010; 16: 1707-1717.

22. Inamoto Y, Jagasia M, Wood WA, Pidala J, Palmer J, Khera N *et al.* Investigator feedback about the 2005 NIH diagnostic and scoring criteria for chronic GVHD. Bone Marrow Transplant. 2014; 49: 532-538.

23. Inamoto Y, Pidala J, Chai X, Kurland BF, Weisdorf D, Flowers ME *et al.* Assessment of joint and fascia manifestations in chronic graft-versus-host disease. Arthritis and Rheumatology 2014; 66: 1044-1052.

24. Jagasia M, Giglia J, Chinratanalab W, Dixon S, Chen H, Frangoul H *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.

25. Arora M, Klein JP, Weisdorf DJ, Hassebroek A, Flowers MED, Cutler CS *et al.* Chronic GVHD risk score: a Center for International Blood and Marrow Transplant Research analysis. Blood 2011; 117: 6714-6720.

26. Inamoto Y, Storer BE, Lee SJ, Carpenter PA, Sandmaier BM, Flowers ME *et al.* Failure-free survival after second-line systemic treatment of chronic graft-versus-host disease. Blood 2013; 121: 2340-2346.

27. Jacobs R, Tran U, Chen H, Kassim A, Engelhardt BG, Greer JP *et al.* Prevalence and risk factors associated with development of ocular GVHD defined by NIH consensus criteria. Bone Marrow Transplant. 2012; 47: 1470-1473.

28. Ogawa Y, Kim SK, Dana R, Clayton J, Jain S, Rosenblatt MI *et al.* International Chronic Ocular Graft-vs-Host-Disease (GVHD) Consensus Group: proposed diagnostic criteria for chronic GVHD (Part I). Science Reporter 2013; 3: 3419.

29. Zantomio D, Grigg AP, Macgregor L, Panek-Hudson Y, Szer J, Ayton R. Female genital tract graft-versus-host disease: incidence, risk factors and recommendations for management. Bone Marrow Transplant. 2006; 38: 567-572.

30. Marks C, Stadler M, Hausermann P, Wolff D, Buchholz S, Stary G *et al.* German-Austrian-Swiss Consensus Conference on clinical practice in chronic graft-versus-host disease (GVHD): guidance for supportive therapy of chronic cutaneous and musculoskeletal GVHD. Br. J. Dermatol. 2011; 165: 18-29.

31. Smith Knutsson E, Björk Y, Broman AK, Helström L, Levin Jakobsen AM, Nilsson O *et al.* Genital chronic graft-versus-host disease in females: a cross-sectional study. Biol Blood Marrow Transplant 2014: [Epub ahead of print 2014 Mar 2]-doi: 10.1016/j.bbmt.2014.02.016.

32. Hirsch P, Leclerc M, Rybojad M, Petropoulou AD, Robin M, Ribaud P *et al.* Female genital chronic graft-versus-host disease: importance of early diagnosis to avoid severe complications. Transplantation 2012; 93: 1265-1269.

33. Spinelli S, Chiodi S, Costantini S, van Lint MT, Raiola AM, Ravera GB *et al.* Female genital tract graft-versus-host disease following allogeneic bone marrow transplantation. Haematologica 2003; 88: 1163-1168.

34. Shanis D, Merideth M, Pulanic TK, Savani BN, Battiwalla M, Stratton P. Female long-term survivors after allogeneic hematopoietic stem cell transplantation: evaluation and management (Review). Semin. Hematol. 2012; 49: 83-93.

35. Mueller SM, Haeusermann P, Rovo A, Halter JP, Passweg J, Itin P *et al.* Genital chronic GVHD in men after hematopoietic stem cell transplantation: a single-center cross-sectional analysis of 155 patients. Biol. Blood. Marrow. Transplant. 2013; 19: 1574-1580.

36. Marks CL, Leiber C, Bertz H, Finke J, Marks R. Genital mucosa in men is a target of chronic graft-versus-host disease with serious impact on quality of life. Bone Marrow Transplant. 2011; 46 (Supplement 1): S66-Abstract #O390.

37. Brook OR, Mullan CP, Mendiratta-Lala M, Joyce R, Sheiman R, Brook A *et al.* Pancreatic atrophy in patients with chronic graft-versus-host disease. Abdominal Imaging 2014; 39: 342-347.

38. Williams KM, Chien JW, Gladwin MT, Pavletic SZ. Bronchiolitis obliterans after allogeneic hematopoietic stem cell transplantation. JAMA 2009; 302: 306-314.

39. Hildebrandt GC, Fazekas T, Lawitschka A, Bertz H, Greinix H, Halter J *et al.* Diagnosis and treatment of pulmonary chronic GVHD: report from the consensus conference on clinical practice in chronic GVHD (Review). Bone Marrow Transplant. 2011; 46: 1283-1295.

40. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. American Journal of Respiratory & Critical Care Medicine 1999; 159: 179-187.

41. Gálban CJ, Han MK, Boes JL, Chughtai KA, Meyer CR, Johnson TD *et al.* Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. Nat. Med. 2012; 18: 1711-1715.

42. Palmer J, Williams K, Inamoto Y, Chai X, Martin PJ, Tomas LS *et al.* Pulmonary symptoms measured by the National Institutes of Health lung score predict overall survival, nonrelapse mortality, and patient-reported outcomes in chronic graft-versus-host disease. Biol Blood Marrow Transplant 2014; 20: 337-344.

43. Pidala J, Chai X, Martin P, Inamoto Y, Cutler C, Palmer J *et al.* Hand grip strength and 2-minute walk test in chronic graft-versus-host disease assessment: analysis from the Chronic GVHD Consortium. Biol. Blood. Marrow. Transplant. 2013; 19: 967-972.

44. Aki SZ, Inamoto Y, Storer BE, Carpenter PA, Lee SJ, Martin P *et al.* Confounding factors affecting the National Institutes of Health (NIH) chronic GVHD organ-specific score and global severity. Biol Blood Marrow Transplant 2014; 20(2, Supplement): S265-S266.

45. Jacobsohn DA, Arora M, Klein JP, Hassebroek A, Flowers ME, Cutler CS *et al.* Risk factors associated with increased nonrelapse mortality and with poor overall survival in children with chronic graft-versus-host disease. Blood 2011; 118: 4472-4479.

46. Lee SJ, Klein JP, Barrett AJ, Ringden O, Antin JH, Cahn JY *et al.* Severity of chronic graft-versus-host disease: association with treatment-related mortality and relapse. Blood 2002; 100: 406-414.

47. Pavletic SZ, Carter SL, Kernan NA, Henslee-Downey J, Mendizabal AM, Papadopoulos E *et al.* Influence of T-cell depletion on chronic graft-versus-host disease: results of a multicenter randomized trial in unrelated marrow donor transplantation. Blood 2005; 106: 3308-3313.

48. Akpek G, Lee SJ, Flowers ME, Pavletic SZ, Arora M, Lee S *et al.* Performance of a new clinical grading system for chronic graft-versus-host disease: a multi-center study. Blood 2003; 102: 802-809.

49. Stewart BL, Storer B, Storek J, Deeg HJ, Storb R, Hansen JA *et al.* Duration of immunosuppressive treatment for chronic graft-versus-host disease. Blood 2004; 104: 3501-3506.

50. Sullivan KM, Witherspoon RP, Storb R, Weiden P, Flournoy N, Dahlberg S *et al.* Prednisone and azathioprine compared with prednisone and placebo for treatment of chronic graft-versus-host disease: Prognostic influence of prolonged thrombocytopenia after allogeneic marrow transplantation. Blood 1988; 72: 546-554.

51. Wingard JR, Piantadosi S, Vogelsang GB, Farmer ER, Jabs DA, Levin LS *et al.* Predictors of death from chronic graft versus host disease after bone marrow transplantation. Blood 1989; 74: 1428-1435.

Table 1. Signs and symptoms of chronic GVHD

ORGAN OR SITE	**DIAGNOSTIC** *(Sufficient to establish the diagnosis of chronic GVHD)*	**DISTINCTIVE** *(Seen in chronic GVHD, but insufficient alone to establish a diagnosis of chronic GVHD)*	**OTHER FEATURES OR UNCLASSIFIED ENTITIES***	**COMMON** *(Seen with both acute and chronic GVHD)*
Skin	• Poikiloderma • Lichen planus-like features • Sclerotic features • Morphea-like features • Lichen sclerosus-like features	• Depigmentation • Papulosquamous lesions	• Sweat impairment • Ichthyosis • Keratosis pilaris • Hypopigmentation • Hyperpigmentation	• Erythema • Maculopapular rash • Pruritus
Nails		• Dystrophy • Longitudinal ridging, splitting or brittle features • Onycholysis • Pterygium unguis • Nail loss** (usually symmetric, affects most nails)		
Scalp and Body Hair		• New onset of scarring or non-scarring scalp alopecia, (after recovery from chemoradiotherapy) • Scaling.	• Thinning scalp hair, typically patchy, coarse or dull (not explained by endocrine or other causes), • Premature gray hair	
Mouth	• Lichen planus-like changes	• Xerostomia • Mucoceles • Mucosal atrophy • Ulcers and Pseudomembranes **		• Gingivitis • Mucositis • Erythema • Pain
Eyes		• New onset dry, gritty, or painful eyes • Cicatricial conjunctivitis • Keratoconjunctivitis sicca • Confluent areas of punctate keratopathy	• Photophobia • Periorbital hyperpigmentation • Blepharitis (erythema of the eye lids with edema)	
Genitalia *Females* *Males*	• Lichen planus-like features • Lichen sclerosus-like features • Vaginal scarring or clitoral/labial agglutination • Phymosis or urethral/meatus scarring or stenosis	• Erosions** • Fissures** • Ulcers**		

ORGAN OR SITE	**DIAGNOSTIC** *(Sufficient to establish the diagnosis of chronic GVHD)*	**DISTINCTIVE** *(Seen in chronic, but insufficient alone to establish a diagnosis of chronic GVHD)*	**OTHER FEATURES OR UNCLASSIFIED ENTITIES**[†]	**COMMON** *(Seen with both acute and chronic GVHD)*
GI Tract	• Esophageal web • Strictures or stenosis in the upper to mid third of the esophagus**		• Exocrine pancreatic insufficiency	• Anorexia • Nausea • Vomiting • Diarrhea • Weight loss • Failure to thrive (infants and children
Liver				• Total bilirubin, alkaline phosphatase > 2 x upper limit of normal • ALT> 2x upper limit of normal
Lung	• Bronchiolitis obliterans diagnosed with lung biopsy • Bronchiolitis obliterans syndrome	•Air trapping and bronchiectasis on chest CT	[†] Cryptogenic organizing pneumonia (COP) [†]Restrictive lung disease	
Muscles, Fascia, Joints	• Fasciitis • Joint stiffness or contractures secondary to sclerosis	• Myositis or polymyositis[††]	• Edema • Muscle cramps • Arthralgia or arthritis	
Hematopoietic and Immune			• Thrombocytopenia • Eosinophilia • Lymphopenia • Hypo- or hyper-gammaglobulinemia • Autoantibodies (AIHA, ITP) • Raynaud's phenomenon	
Other			• Pericardial or pleural effusions • Ascites • Peripheral neuropathy • Nephrotic syndrome • Myasthenia gravis • Cardiac conduction abnormality or cardiomyopathy	

*Can be acknowledged as part of the chronic GVHD symptomatology if diagnosis is confirmed
**In all cases, infection, drug effect, malignancy or other causes must be excluded.
[†] Pulmonary entities under investigation or unclassified.

†† Diagnosis of chronic GVHD requires biopsy

Abbreviation: ALT (alanine aminotransferase); AST (aspartate aminotransferase); PFTs (pulmonary function tests); AIHA (autoimmune hemolytic anemia); ITP (idiopathic thrombocytopenic purpura).

Table 2 - NIH Global Severity of Chronic GVHD

Mild chronic GVHD

1or 2 organs involved *plus*
Score in involved organs 1 *plus*
Lung score 0

Moderate chronic GVHD

3 or more organs involved *plus*
Score of 1 in each organ

OR

At least 1 organ (not lung) with a score of 2

OR

Lung score 1

Severe chronic GVHD

At least 1 organ with a score of 3

OR

Lung score of 2 or 3

Key points:

1. In skin: higher of the two scores to be used for calculating global severity.
2. In lung: FEV1 is used instead of clinical score for calculating global severity.
3. If the entire abnormality in an organ is annotated as unequivocally explained by a non-GVHD documented cause, that organ severity score will be downgraded to a zero score for calculation of the global severity only.
4. If the abnormality in an organ is attributed to multifactorial causes (GVHD plus other causes) the scored organ will be used for calculation of the global severity regardless of the contributing causes (no downgrading of organ severity score).

Figure 1. Organ Scoring of Chronic GVHD

	SCORE 0	SCORE 1	SCORE 2	SCORE 3
PERFORMANCE SCORE: ☐ **KPS ECOG LPS**	☐ Asymptomatic and fully active (ECOG 0; KPS or LPS 100%)	☐ Symptomatic, fully ambulatory, restricted only in physically strenuous activity (ECOG 1, KPS or LPS 80-90%)	☐ Symptomatic, ambulatory, capabl of self-care, >50% of waking hours ou of bed (ECOG 2, KPS or LPS 60-70%)	☐ Symptomatic, limited self-care, >50% of waking hours in bed (ECOG 3-4, KPS or LPS <60%)
SKIN† **SCORE % BSA** ☐ *GVHD features to be scored by BSA:* **Check all that applies:** ☐ Maculopapular rash/erythema ☐ Lichen planus-like features ☐ Sclerotic features ☐ Papulosquamous lesions or ichthyosis ☐ Keratosis pilaris-like GVHD	☐ No BSA involved	☐ 1-18% BSA	☐ 19-50% BSA	☐ >50% BSA
SKIN FEATURES SCORE:	☐ No sclerotic features		☐ **S**uperficial sclerotic features "not hidebound" (able to pinch)	**Check all that applies:** ☐ Deep sclerotic features ☐ "Hidebound" (unable to pinch) ☐ Impaired mobility ☐ Ulceration

Other skin GVHD features (NOT scored by BSA)
Check all that applies:
☐ Hyperpigmentation
☐ Hypopigmentation
☐ Poikiloderma
☐ Severe or generalized pruritus
☐ Hair involvement
☐ Nail involvement
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____

MOUTH *Lichen planus-like features present:* ☐ **Yes** ☐ **No**	☐ No symptoms	☐ Mild symptoms **with** disease signs but not limiting oral intake significantly	☐ Moderate symptoms with disease signs **with** partial limitation of oral intake	☐ Severe symptoms with disease signs on examination **with** major limitation of oral intake

☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____

Figure 1. Organ Scoring of Chronic GVHD (continued)

	SCORE 0	SCORE 1	SCORE 2	SCORE 3
EYES *Keratoconjunctivitis sicca (KCS) confirmed by Ophthalmologist:* ☐ **Yes** ☐ **No** ☐ **Not examined**	☐ No symptoms	☐ Mild dry eye symptoms not affecting ADL (requirement of lubricant eye drops ≤ 3 x per day)	☐ Moderate dry eye symptoms partially affecting ADL (requiring lubricant eye drops > 3 x per day or punctal plugs), **WITHOUT** new vision impairment due to KCS	☐ Severe dry eye symptoms significantly affecting ADL (special eyeware to relieve pain) **OR** unable to work because of ocular symptoms **OR** loss of vision due to KCS
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*				
GI Tract ***Check all that applies:*** ☐ Esophageal web/ proximal stricture or ring ☐ Dysphagia ☐ Anorexia ☐ Nausea ☐ Vomiting ☐ Diarrhea ☐ Weight loss* ☐ Failure to thrive	☐ No symptoms	☐ Symptoms without significant weight loss* (<5%)	☐ Symptoms associated with mild to moderate weight loss* (5-15%) **OR** moderate diarrhea without significant interference of daily living	☐ Symptoms associated with significant weight loss* >15%, requires nutritional supplement for most calorie needs **OR** esophageal dilation **OR** severe diarrhea with significant interference of daily living
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____				
LIVER	☐ Normal total bilirubin and ALT or AP < 3 x ULN	☐ Normal total bilirubin with ALT ≥3 to 5 x ULN or AP ≥ 3 x ULN	☐ Elevated total bilirubin but ≤3 mg/dL or ALT > 5 ULN	☐ Elevated total bilirubin > 3 mg/dL
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*				
LUNGS**				
Symptoms score:	☐ No symptoms	☐ Mild symptoms (shortness of breath after climbing one flight of steps)	☐ Moderate symptoms (shortness of breath after walking on flat ground)	☐ Severe symptoms (shortness of breath at rest; requiring 0_2)
Lung score: FEV1 ☐	☐ FEV1≥80%	☐ FEV1 60-79	☐ FEV1 40-59%	☐ FEV1 ≤39%
Pulmonary function tests ☐ Not performed				
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____				

Figure 1. Organ scoring of chronic GVHD (continued)

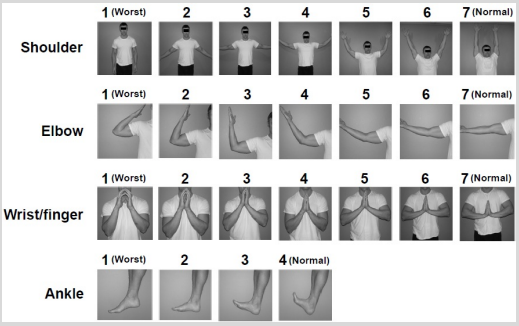
	SCORE 0	SCORE 1	SCORE 2	SCORE 3
JOINTS AND FASCIA P-ROM score *(see below)* Shoulder (1-7):___ Elbow (1-7): ____ Wrist/finger (1-7):___ Ankle (1-4):___	☐ No symptoms	☐ Mild tightness of arms or legs, normal or mild decreased range of motion (ROM) **AND** not affecting ADL	☐ Tightness of arms or legs **OR** joint contractures, erythema thought due to fasciitis, moderate decrease ROM **AND** mild to moderate limitation of ADL	☐ Contractures **WITH** significant decrease of ROM ***AND*** significant limitation of ADL (unable to tie shoes, button shirts, dress self etc.)
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____				
GENITAL TRACT (*See Supplemental figure*‡) **Check all that applies** ☐ Not examined *Currently sexually active* ☐ Yes ☐ No	☐ No signs	☐ Mild signs‡ and females with or without discomfort on exam	☐ Moderate signs‡ and may have symptoms* with discomfort on exam	☐ Severe signs‡ with or without symptoms
☐ *Abnormality present but explained entirely by non-GVHD documented cause (specify):*_____				

Other indicators, clinical features or complications related to chronic GVHD (check all that apply and assign a score to its severity (0-3) based on its functional impact where applicable none – 0,mild -1, moderate -2, severe – 3)

☐ Ascites (serositis)___	☐ Myasthenia Gravis___	
☐ Pericardial Effusion___	☐ Peripheral Neuropathy___	☐ Eosinophilia > 500µl___
☐ Pleural Effusion(s)___	☐ Polymyositis___	☐ Platelets <100,000/µl ___
☐ Nephrotic syndrome___	☐ Weight loss* without GI symptoms ___	☐ Others (specify):_____

Overall GVHD Severity *(Opinion of the evaluator)* ☐ **No GVHD** ☐ **Mild** ☐ **Moderate** ☐ **Severe**

Photographic Range of Motion (P-ROM)



† Skin scoring should use both percentage of BSA involved by disease signs and the cutaneous features scales. When a discrepancy exists between the percentage of total body surface (BSA) score and the skin feature score, OR if superficial sclerotic features are present (Score 2), but there is impaired mobility or ulceration (Score 3), the higher level should be used for the final skin scoring.

* Weight loss within 3 months.

**Lung scoring should be performed using both the symptoms and FEV1 scores whenever possible. FEV1 should be used in the final lung scoring where there is discrepancy between symptoms and FEV1 scores.

Abbreviations: ECOG (Eastern Cooperative Oncology Group), KPS (Karnofsky Performance Status), LPS (Lansky Performance Status); BSA (body surface area); ADL (activities of daily living); LFTs (liver function tests); AP (alkaline phosphatase); ALT (alanine aminotransferase); NUL (normal upper limit).

‡ To be completed by specialist or trained medical providers (see Supplemental Figure).

Supplement Figure 1 – Genital Tract Chronic Graft-versus-Host Assessment and Scoring Form

Name:_____ **Date of birth:** _____**Assessment date:** _____

	SCORE 0	SCORE 1	SCORE 2	SCORE 3
GENITAL TRACT (male or female)	☐ No signs	☐ Mild signs and females may have symptoms* WITH discomfort on exam	☐ Moderate signs and may have symptoms* with discomfort on exam	☐ Severe signs with or without symptoms *

Currently sexually active:
☐ Yes ☐ No

Check all signs that applies:

☐ Lichen planus-like features

☐ Lichen sclerosis-like features

☐ Vaginal scarring (female)

☐ Clitoral/labial agglutination (female)

☐ Labial resorption (female)

☐ Erosions

☐ Fissures

☐ Ulcers

☐ Phimosis (male)

☐ Urethral meatus scarring/ stenosis (male)

☐ *Abnormality present but NOT thought to represent GVHD (specify cause):*_____
☐ *Abnormality thought to represent GVHD PLUS other causes(specify cause):*_____

**Genital symptoms are not specific to cGVHD and can represent premature gonadal failure or genital tract infection*.

If a gynecologist is unavailable, external examination may be performed to determine "discomfort on exam" as follows:

a) Spread the labia majora to inspect the vulva for the above signs. Touch the vestibular gland openings (Skene's and Bartholin's), labia minora and majora gently with a qtip. Vulvar pain elicited by the gentle touch of a qtip is classified as discomfort on examination. Palpate the vaginal walls with a single digit to detect bands, shortening, narrowing or other signs of vaginal scarring.
b) If the woman is sexually active, determine whether qtip palpation or gentle palpation of scarred ridges elicits pain similar to that which the woman experiences during intercourse.

Female genitalia: Severity of signs:

1) Mild (any of the following); erythema on vulvar mucosal surfaces, vulvar lichen-planus or vulvar lichen-sclerosis.
2) Moderate (any of the following); erosive inflammatory changes of the vulvar mucosa, fissures in vulvar folds
3) Severe (any of the following); labial fusion, clitoral hood agglutination, fibrinous vaginal adhesions, circumferential fibrous vaginal banding, vaginal shortening, synechia, dense sclerotic changes, and complete vaginal stenosis.