

REVIEW

Chronic graft-versus-host disease: implications of the National Institutes of Health consensus development project on criteria for clinical trials

SZ Pavletic¹, SJ Lee², G Socie³ and G Vogelsang⁴

¹Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; ²Division of Clinical Research, Fred Hutchinson Cancer Research Center, University of Washington School of Medicine, Seattle, WA, USA; ³Department of Hematology-Transplantation, AP-HP Hospital Saint Louis, Service d'Hématologie Greffe, Paris, France and ⁴The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Chronic graft-versus-host disease (GVHD) has been a difficult problem to address and clinical research in this area lags behind other innovations in hematopoietic stem cell transplantation (HCT). Recently the international transplant community has focused more on chronic GVHD. This new focus is well represented by the development of the National Institutes of Health sponsored chronic GVHD consensus project, which has unified the transplant community's approach to chronic GVHD through the activities of focused working groups. From December 2005 through May 2006, a series of consensus documents have been published addressing the areas of diagnosis and staging, histopathology, strategies for the development and validation of biomarkers, response criteria, ancillary therapy and supportive care and the design of clinical trials. This paper summarizes and discusses these reports, focusing specifically on diagnosis and scoring and response criteria. Although these documents represent a huge effort by the research community, they must be prospectively implemented and validated. These new criteria should advance the standards and uniformity of chronic GVHD clinical research. The ultimate success of this project is dependent on whether these recommendations move the field forward. This is an opportunity for the transplant community to unite and make a significant impact in chronic GVHD.

Bone Marrow Transplantation (2006) 38, 645–651. doi:10.1038/sj.bmt.1705490; published online 18 September 2006

Keywords: chronic graft-versus-host disease; consensus; guidelines; stem cell transplantation

Introduction

Chronic graft-versus-host disease (GVHD) has been a difficult problem to address and clinical research seeking to better understand chronic GVHD lags behind other innovations in hematopoietic stem cell transplantation (HCT). The original descriptions and the staging system were based on small numbers of patients in a pre-cyclosporine era.¹ The issues which prevented early work in chronic GVHD – patients have returned home when the disease begins, no consistent laboratory findings, profound immune incompetence resulting in frequent severe infections, and variable presentation and organ involvement – have continued to hinder study and treatment of chronic GVHD. Although different centers and groups have tried to tackle specific aspects-suggesting new staging systems and response criteria, it was obvious that a concerted effort looking at the total disorder was needed.² At the same time chronic GVHD receives little attention at the national meetings, and there is no US Food and Drug Administration (FDA)-approved medication for chronic GVHD.

The Chronic GVHD Consensus Project was sponsored by the intramural and extramural programs of the National Cancer Institute, National Heart Lung and Blood Institute, and National Institute for Allergy and Infectious Disease, the NIH Director's Office, the Health Resources and Services Administration, and the Department of Defense, and the Naval Medical Research Center. The FDA representatives were involved as a major goal was the development of new therapies for chronic GVHD. The first planning meeting was held in June 2004 and included transplant investigators with a focus on chronic GVHD, consultants, and representatives of government agencies. Six working groups were formed: diagnosis and staging of chronic GVHD, histopathology, biomarkers, response criteria, supportive care, and design of clinical trials. These groups were inclusive to ensure their work reflected consensus opinion. It was clear from the outset that there were insufficient data to provide evidence-based recommendations. Thus, the reports of these groups represent the best expert opinion.

In June 2005, the NIH Consensus Meeting was held in Washington, DC Members of the European Blood and

Correspondence: Dr SZ Pavletic, Graft-Versus-Host and Autoimmunity Unit, Experimental Transplantation and Immunology Branch, National Cancer Institute, 9000 Rockville Pike, CRC 3-3330, Bethesda, MD 20892-1203, USA.

E-mail: pavletis@mail.nih.gov

Received 2 June 2006; revised 31 July 2006; accepted 2 August 2006; published online 18 September 2006

Marrow Transplant Group (EBMT), American Society for Blood and Marrow Transplantation (ASBMT), Center for International Blood and Marrow Transplant Research (CIBMTR), industry representatives, NIH scientists and patient representatives were invited. The meeting featured a series of panel discussions to present the working group reports and lectures focusing on issues critical for future research directions in the field. Detailed reports from each of six working groups have been published in *Biology of Blood and Marrow Transplantation* but no concise report has been published so far summarizing the project.³⁻⁸ This paper provides a summary statement with some background and commentaries to put these publications in context (Table 1). It focuses specifically on recommendations in diagnosis and scoring and response criteria because these reports contain material likely to set the standards for anyone contemplating clinical trials in chronic GVHD and contribute to this project by using the proposed tools.

The ultimate success of this project is dependent on whether these recommendations move the field forward. This is an opportunity for the transplant community to unite and make a significant impact in this disease. None of the participants in this project believe that their reports are perfect and their work is done. However, all agree that this is a communal first step and share in the hope that the spirit, which has produced this effort will continue to refine these instruments and use them to conduct studies that improve our understanding of chronic GVHD.

Diagnosis of chronic GVHD

The Diagnosis and Scoring paper offers an updated conceptualization of chronic GVHD that is based on the specificity of signs rather than the traditional criterion of time of onset since transplantation (more or less than 100 days).³ Thus, the diagnosis of chronic GVHD requires at least one *diagnostic* sign that is so classic for chronic GVHD that no additional testing is required (examples include poikiloderma, sclerotic skin features, oral lichen-type changes) or at least one *distinctive* sign that is highly suggestive (for example nail dystrophy, vitiligo-like depigmentation or bronchiolitis obliterans diagnosis based only on pulmonary function tests and computerized tomography findings) but requires laboratory or histopathologic confirmation in the same or other organ (Figure 1). A biopsy read as 'consistent with' or 'unequivocal' GVHD will be considered sufficient to support the diagnosis of chronic GVHD if accompanied by at least one distinctive clinical manifestation.⁴ As always, appropriate studies should be performed to rule out other potential diagnoses such as infection, drug toxicity etc.

In addition to setting diagnostic criteria for chronic GVHD, the consensus conference realized that it would need to provide a definition of acute GVHD so that the two syndromes could be differentiated. Some signs and symptoms are common to both chronic and acute GVHD (erythema, maculopapular rash, nausea, vomiting or

Table 1 Documents published by NIH consensus development project on criteria for clinical trials in chronic GVHD

Document	Key definitions and recommendations included
Diagnosis and staging ³	Minimal clinical diagnostic criteria Clinical distinction between acute and chronic GVHD New organ severity scoring system New global scoring system Indications for systemic therapy Strategies for developing better predictors of TRM
Histopathology ⁴	Minimal diagnostic criteria for active GVHD Biopsy features suggestive of chronic GVHD Standardized terminology for reporting results List of clinical data that should accompany biopsy Organ-specific criteria for adequate tissue sampling Standardized research forms for histology reporting (at – http://www.asbmt.org/cGVHD_Guidelines.htm)
Biomarkers ⁵	Rationale for developing biomarkers in chronic GVHD Definition of biomarkers and potential applications Proposal for a pathophysiology-based classification Current data in search for biomarkers in chronic GVHD Methodological considerations in identification and validation of biomarkers in chronic GVHD
Response criteria ⁶	Proposal of quantifiable set of chronic GVHD measures for outpatient use by transplant and non-transplant providers and in adult and pediatric settings Introduction of ancillary objective and patient reported measures of function, performance and quality-of-life Provisional criteria and guidelines for determining partial response and progression Guidelines for use of response assessments in trials
Ancillary and supportive care ⁷	Evidence based rated organ-specific guidelines for ancillary and supportive care including: skin, mouth, eyes, vulvar-vaginal, gastrointestinal-liver, lungs, hematopoietic, neurologic, immunologic-infection prevention, musculoskeletal, psychosocial (dispensary guidelines at – http://www.asbmt.org/GVHDForms.htm) Clinical monitoring guidelines in chronic GVHD patients Need for multidisciplinary team approach
Design of clinical trials ⁸	Approaches and definitions to ensure consistent and interpretable results of clinical studies designed to assess interventions for treatment of chronic GVHD

Diagnostic versus Non-Diagnostic 'Other' Features

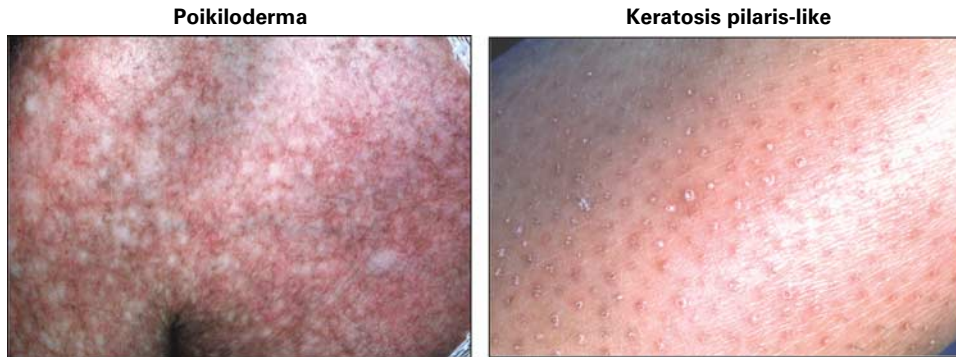


Figure 1 *Poikiloderma* – an example of a 'Diagnostic' chronic GVHD sign that does not require biopsy or other confirmation for establishing the clinical diagnosis of chronic GVHD. *Keratosis pilaris* is in contrast an example of 'other' manifestations, clinical signs in that group can not be used to establish the diagnosis of chronic GVHD but can be interpreted as part of the chronic GVHD symptomatology if the diagnosis of chronic GVHD is confirmed.³ Photographs provided kindly by Edward Cowen and Maria Turner, Dermatology Branch, NCI.

diarrhea, or elevated liver function tests) and thus cannot be used to distinguish the two. Instead, the Diagnosis and Scoring document delineates two main categories of GVHD, each with two subcategories (Table 2). The broad category of acute GVHD includes *classic* acute GVHD (maculopapular erythematous rash, gastrointestinal symptoms, or cholestatic hepatitis) occurring within 100 days post-transplant or donor leukocyte infusion (DLI) while *persistent, recurrent or late acute GVHD* (usually seen after withdrawal of immunosuppression) occurs beyond 100 days of transplantation or DLI. The arbitrary day 100 distinction is retained in separating of these two acute GVHD categories to facilitate reporting in clinical trials. Both acute GVHD sub entities should occur without presence of diagnostic or distinctive chronic GVHD manifestations. A second broad GVHD category encompasses *classic* chronic GVHD consisting only of manifestations that can be ascribed to chronic GVHD; and *acute and chronic overlap syndrome* in which features of both acute and chronic GVHD appear together. In the absence of histological or clinical signs or symptoms characteristic of chronic GVHD, persistence, recurrence or new onset of characteristic skin, gastro intestinal (GI) tract or liver abnormalities should be classified as acute GVHD regardless of the time after transplantation. With appropriate stratifications, patients with persistent, recurrent or late acute GVHD or overlap syndrome can be included in clinical trials with patients who have chronic GVHD.

Scoring of chronic GVHD

A clinical categorical system (0–3) is recommended for scoring of individual organs that describes the severity for each affected organ/site at any given time taking functional impact into account. This system is applied only after the diagnosis of chronic GVHD is confirmed. Organs/sites assessed for scoring include: skin, mouth, eyes, GI tract, liver, lungs, joints and fascia, and genital tract. In general, a score of zero means no manifestations/symptoms, a score of 1 means no significant impairment of function

Table 2 GVHD categories

Category	Time of manifestation after HCT or DLI	Presence of GVHD features	
		Acute	Chronic
<i>Acute GVHD</i>			
Classic acute GVHD	≤ 100 days	Yes	No
Persistent, recurrent or late onset acute GVHD	> 100 days	Yes	No
<i>Chronic GVHD</i>			
Classic chronic	No time limit	No	Yes
Overlap syndrome	No time limit	Yes	Yes

or activities of daily living (ADL), a score of 2 means significant impairment of ADL but no major disability and a score of 3 indicates significant impairment of ADL with major disability. The scoring can be easily conducted in the clinic and the only mandated laboratory tests for its completion are liver function tests. Utilization of pulmonary function tests for FEV1 and DLCO assessments is strongly encouraged for lung scoring although this can be done also by assessing the symptoms of shortness of breath and need for supplemental oxygen. An example of the 0–3 organ severity scoring is shown for skin (Table 3).

A global assessment of severity (none, mild, moderate, severe) is derived by combining organ-specific scores and was intended to replace the current 'limited-extensive' scoring system after appropriate validation.

Indications for systemic therapy

Symptomatic mild chronic GVHD may often be treated with local therapies alone (e.g., topical steroids to the skin or cyclosporine eye drops). However, systemic therapy should be considered for patients who meet criteria for moderate-severe global severity (involvement of three or more organs or with a score of 2 or greater in any single

Table 3 Chronic GVHD organ severity scoring example in skin

Chronic GVHD scoring – skin			
0	1	2	3
No symptoms	≤18% BSA with disease signs but <i>no</i> sclerotic features	19–50% BSA <i>or</i> involvement with superficial sclerotic features ‘not hidebound’ (able to pinch)	> 50% BSA <i>or</i> deep sclerotic features ‘hidebound’ (unable to pinch) <i>or</i> impaired mobility, ulceration or severe pruritus

organ, or any lung involvement). Good medical practice and judgment dictate flexibility in this recommendation. Some experts incorporate the presence or absence of published high-risk features (e.g., thrombocytopenia or progressive onset) and the underlying reason for transplantation (e.g., malignant versus nonmalignant disease) or current co-morbid conditions (e.g., infection) into the decision of whether or not to treat with systemic immunosuppression. Appropriate early intervention with effective systemic therapy may prevent progression to severe chronic GVHD. Patients with chronic GVHD, are immunocompromised and should receive infection-prevention measures as outlined in the Ancillary Therapy and Supportive Care Working Group document.⁷

Prognostic indicators

Chronic GVHD is a major cause of late transplant-related mortality (TRM) after allogeneic HCT.⁹ Across studies, thrombocytopenia (platelet count <100 000/ μ l) and progressive onset of chronic GVHD from acute GVHD consistently predict an increased risk of TRM.¹⁰ Validation of risk factors for TRM in patients with chronic GVHD should be a major goal of future research so that patients with the poorest prognoses will be included in clinical trials of systemic therapies aimed at changing the natural history of chronic GVHD. Conversely, patients judged at low risk for TRM might be preferentially enrolled in studies of new topical or organ-specific therapies.

Measuring response in chronic GVHD clinical trials

There are no validated, organ-specific or overall response criteria for chronic GVHD. This lack of standardized criteria for evaluation of therapeutic response is a major obstacle in clinical trials and care of patients with chronic GVHD. Current definitions of response are based on subjective assessments of global complete (disappearance of all symptoms) or partial response (50% reduction in symptoms) based on non-reproducible physician assessments. There is a considerable variability in tools used from one study to the next (review by G Akpek available at: <http://www.asbmt.org/GVHDForms.htm>). There has been a trend towards introducing ‘surrogate’ measures of benefit in chronic GVHD trials such as the ability to taper steroids. However, this end point is controlled by the physician and of debatable value unless the taper is conducted according to a fixed predefined algorithm or the trial is blinded.

Chronic GVHD is a disease with variable manifestations and response to therapy. Some chronic GVHD manifestations may be irreversible even with appropriate therapy. Thus, the consensus document proposes a variety of measures to quantify disease activity in Phase I–III trials evaluating chronic GVHD therapy. The tools were based on existing instruments or strategies used in chronic GVHD and in other fields of medicine.^{11–14} The recommended measures address research requirements and do not necessarily reflect practices that might apply to routine patient care. Measures should be made at 3-month intervals and whenever a major change is made in treatment. Blinding of treatment or evaluators is optimal.

Taking into account many considerations, including the variable expertise of the evaluator, the heterogeneity of chronic GVHD manifestations and the specific aims of clinical trials, the consensus conference recommends the following assessments.

Chronic GVHD-specific core measures include:

- Clinician or patient-assessed signs and symptoms (specific to chronic GVHD) (Table 4).
- Chronic GVHD symptom scale by Lee *et al.* (a partially validated patient self reported scale).¹⁵
- The clinician- or patient-reported global rating scales (4-point, 7-point and 11-point physician or patient reported ratings of chronic GVHD severity or its trajectory) are included to quantify qualitative components of clinical assessments. These tools are frequently used in other areas of medicine for quantification of responses.^{15–17}

Chronic GVHD-nonspecific ancillary measures (optional, secondary endpoints):

- Measurement of functional performance, grip strength and 2-min walk time.^{18–20}
- Human Activity Profile (HAP) questionnaire (a patient-reported measure of functional capacity).²¹
- Clinician-assessed Karnofsky performance status (included for historical reasons, not considered sufficiently sensitive for response assessment).
- The SF-36 version two questionnaire and FACT-BMT for quality-of-life assessments.^{22–24}

Practical forms to be used for clinician and patient reported chronic GVHD assessments are provided at <http://www.asbmt.org/GvHDForms> (Forms A and B). The same website includes the complete PowerPoint educational manual that can be downloaded with precise instructions how to use NIH consensus recommended response criteria and the picture dictionary of skin and oral chronic GVHD manifestations.

Table 4 Clinician assessed and patient reported chronic GVHD signs and symptoms

<i>Organ</i>	<i>Sign or symptom item</i>	<i>Measure</i>
Skin	Erythematous rash of any sort	% Body surface area
	Movable sclerosis	0–100% for each item by using rules of nines
	Nonmoveable sclerosis, or subcutaneous sclerosis/fasciitis	
	Ulcers	Largest dimension (cm) of the largest ulcer
	Pruritus or itching	0–10 patient reported intensity scale
Eyes	Bilateral Schirmer's tear test scores without anesthesia ^a	Mean of both eyes (mm)
	Chief ocular complaint at the time of the visit	0–10 patient reported intensity scale
Mouth	Erythema	Total score 0–15
	Lichen-type Hyperkeratosis	
	Ulcerations	0–10 patient reported intensity scale
	Mucoceles	
	Symptoms of oral pain, dryness, sensitivity	
Hematology	Platelet count	Number/ μ l
	Eosinophils	Percent
Gastrointestinal	Upper GI symptoms	0–3
	Esophageal symptoms	0–3
	Diarrhea	0–3
Liver	Total serum bilirubin	mg/dl
	ALT, alkaline phosphatase	U/L
Lungs	Bronchiolitis obliterans syndrome	FEV1, DLCO

Vulvar-vaginal symptoms (yes or no) and patient weight are recorded at each visit.

Range of motion of the most affected joints is recorded depending on the availability of a qualified specialist.

^aThe primary proposed objective measure of lacrimal gland function in chronic GVHD is the Schirmer's test (to be performed without anesthesia) for each eye separately, as recommended by the Sjögren's syndrome consensus group.²⁶ Instructions for administration of the Schirmer's test are provided with the instructional manual at: <http://www.asbmt.org/GvHDForms>.

Recommended utilization of staging and response measures

The discussed documents provide a recommended list of core evaluations that should be conducted in chronic GVHD therapeutic or natural history clinical studies. The proposed staging system is a categorical scoring system based on expert opinion meant for quick clinical assessment of the extent and severity of chronic GVHD signs and symptoms at any important baseline assessment point (new diagnosis, new patient, yearly follow-up in the tertiary center, for example). With time and validation it may show utility for stratification in trials, therapy planning, prognostication and perhaps routine clinical practice. While the proposed 0–3 organ scoring system may provide a quick assessment of chronic GVHD, such categorical scales consisting of a composite mixture of symptoms, signs, performance status, activities of daily living and therapeutic interventions were not intended for use as a response criteria measure. Nevertheless, to allow internal validation and future studies, it is recommended that all scales be collected simultaneously. Preliminary evidence suggests these tools will prove feasible and reproducible.²⁵

Future directions

The natural question is what comes next? Although these documents represent a huge effort by the research community, they must be prospectively implemented and validated. The proposed tools should enhance uniformity

of data collection and provide standards for clinical trials. It is hoped that after empirical data are collected, the current response criteria can be condensed down to clinically meaningful assessments proven to be predictive of overall response. This analysis may be facilitated by an outcomes repository for data collected in clinical trials and natural history studies using these instruments.

Organ-site specialists should be engaged to develop methods for more sensitive and objective assessment of specific organs. Future studies should determine the extent to which patient-reported outcomes and functional measures can be used as primary end points in chronic GVHD clinical trials. Improved methods will be needed to distinguish chronic GVHD disease activity from irreversible damage and to develop a chronic GVHD activity index for clinical trials, perhaps through the use of biomarkers.

This project has sparked renewed interest in chronic GVHD research. The recommendations that have emerged from this initiative are being applied in several projects in the United States, Canada, Europe, and Latin America. The broad participation of scientists at all levels in this effort has engaged a whole new generation of investigators. Enthusiasm alone is not sufficient to carry us forward. Both the medical and research communities are severely challenged when chronic GVHD develops far from the transplant centers. Better methods for communication and access to centers with chronic GVHD expertise may improve clinical care and research opportunities. It is clear that innovative and more substantial funding will be

necessary to overcome these obstacles. At the NCI Center for Cancer Research (CCR) an open, interactive website is under development. In the meantime comments can be submitted via the American Society for Blood and Marrow Transplantation chronic GVHD guidelines Web site (http://www.asbmt.org/cGvHD_Guidelines).

It is also important to grasp the broader implications of the NIH consensus effort. Chronic GVHD should be a model for studying new immunomodulating agents. The basic biology of chronic GVHD, which so closely mimics many spontaneously occurring autoimmune disorders and has a profound anti-tumor effect must be deciphered. Understanding these mechanisms would have far reaching clinical implications.

Acknowledgements

The National Institutes of Health (NIH) Consensus Development Project on Criteria for Clinical Trials in Chronic GVHD was supported by the National Cancer Institute, Office of the Director, Cancer Therapy Evaluation Program, Intramural Research Program, and Center for Cancer Research; National Heart Lung and Blood Institute, Division of Blood Diseases and Resources; Office of Rare Diseases, NIH, Office of the Director; National Institute of Allergy and Infectious Disease, Transplantation Immunology Branch; and the Health Resources and Services Administration, Division of Transplantation and the Naval Medical Research Center, CW Bill Young/Department of Defense Marrow Donor Recruitment and Research Program. The authors would also like to acknowledge the following organizations that by their participation made this project possible: American Society for Blood and Marrow Transplantation, European Group for Blood and Marrow Transplantation, Center for International Bone and Marrow Transplant Research, Blood and Marrow Transplant Clinical Trials Network, Canadian Blood and Marrow Transplant Group and Pediatric Blood and Marrow Transplant Consortium. The organizers are also indebted to patients and patient and research advocacy groups who made this process much more meaningful by their engagement. The project group recognizes contributions of numerous colleagues in the field of blood and marrow transplantation, medical specialists and consultants, pharmaceutical industry, and the National Institutes of Health and US Food and Drug Administration professional staff for their intellectual input, dedication, and enthusiasm.

Disclaimer

The opinions expressed here are those of the authors and do not represent the official position of the National Institutes of Health, Food and Drug Administration, or the US Government.

References

- 1 Shulman HM, Sullivan KM, Weiden PL, McDonald GB, Striker GE, Sale GE *et al*. Chronic graft-versus-host syndrome in man. A long-term clinicopathologic study of 20 Seattle patients. *Am J Med* 1980; **69**: 204–217.
- 2 Farag SS. Chronic graft-versus-host disease: where do we go from here? *Bone Marrow Transplant* 2004; **33**: 569–577.

- 3 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.
- 4 Shulman HM, Kleiner D, Lee SJ, Morton T, Pavletic SZ, Farmer E *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.
- 5 Schultz KR, Miklos DB, Fowler D, Cooke K, Shizuru J, Zorn E *et al*. Toward biomarkers for chronic graft-versus-host disease: National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: III. Biomarker Working Group Report. *Biol Blood Marrow Transplant* 2006; **12**: 126–137.
- 6 Pavletic SZ, Martin P, Lee SJ, Mitchell S, Jacobsohn D, Cowen EW *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.
- 7 Couriel D, Carpenter PA, Cutler C, Bolanos-Meade J, Treister NS, Gea-Banacloche J *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.
- 8 Martin PJ, Weisdorf D, Przepiora D, Hirschfeld S, Farrell A, Rizzo JD *et al*. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: VI. Design of Clinical Trials Working Group report. *Biol Blood Marrow Transplant* 2006; **12**: 491–505.
- 9 Socie G, Stone JV, Wingard JR, Weisdorf D, Henslee-Downey PJ, Bredeson C *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.
- 10 Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. *Biol Blood Marrow Transplant* 2003; **9**: 215–233.
- 11 Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B *et al*. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. *Arthritis Rheum* 1993; **36**: 729–740.
- 12 Ward MM, Marx AS, Barry NN. Comparison of the validity and sensitivity to change of 5 activity indices in systemic lupus erythematosus. *J Rheumatol* 2000; **27**: 664–670.
- 13 Sandborn WJ, Feagan BG, Hanauer SB, Lochs H, Lofberg R, Modigliani R *et al*. A review of activity indices and efficacy endpoints for clinical trials of medical therapy in adults with Crohn's disease. *Gastroenterology* 2002; **122**: 512–530.
- 14 Rider LG, Giannini EH, Harris-Love M, Joe G, Isenberg D, Pilkington C *et al*. Defining Clinical Improvement in Adult and Juvenile Myositis. *J Rheumatol* 2003; **30**: 603–617.
- 15 Lee S, Cook EF, Soiffer R, Antin JH. Development and validation of a scale to measure symptoms of chronic graft-versus-host disease. *Biol Blood Marrow Transplant* 2002; **8**: 444–452.
- 16 Cleeland CS, Mendoza TR, Wang XS, Chou C, Harle MT, Morrissey M *et al*. Assessing symptom distress in cancer patients: the M.D. Anderson Symptom Inventory. *Cancer* 2000; **89**: 1634–1646.

- 17 Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 1998; **16**: 139–144.
- 18 Mathiowetz V, Kashman N, Volland G, Weber K, Dowe M, Rogers S. Grip and pinch strength: normative data for adults. *Arch Phys Med Rehabil* 1985; **66**: 69–74.
- 19 Brooks D, Parsons J, Tran D, Jeng B, Gorczyca B, Newton J *et al*. The two-minute walk test as a measure of functional capacity in cardiac surgery patients. *Arch Phys Med Rehabil* 2004; **85**: 1525–1530.
- 20 Eiser N, Willsher D, Dore CJ. Reliability, repeatability and sensitivity to change of externally and self-paced walking tests in COPD patients. *Respir Med* 2003; **97**: 407–414.
- 21 Daughton DM, Fix AJ, Kass I, Bell CW, Patil KD. Maximum oxygen consumption and the ADAPT quality-of-life scale. *Arch Phys Med Rehabil* 1982; **63**: 620–622.
- 22 Ware Jr JE. SF-36 health survey update. *Spine* 2000; **25**: 3130–3139.
- 23 Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30**: 473–483.
- 24 McQuellon RP, Russell GB, Cella DF, Craven BL, Brady M, Bonomi A *et al*. Quality of life measurement in bone marrow transplantation: development of the Functional Assessment of Cancer Therapy-Bone Marrow Transplant (FACT-BMT) scale. *Bone Marrow Transplant* 1997; **19**: 357–368.
- 25 Mitchell S, Reeve B, Cowen E, Fall-Dickson J, Gerber L, Guadagnini JP *et al*. Feasibility and reproducibility of the new NIH consensus criteria to evaluate response in chronic GVHD – a Pilot study. *Blood* 2005; **106** (Suppl 1): abstract #3121.
- 26 Vitali C, Bombardieri S, Jonsson R, Moutsopoulos HM, Alexander EL, Carsons SE *et al*. Classification criteria for Sjogren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. *Ann Rheum Dis* 2002; **61**: 554–558.