Draft May 12, 2014

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

Running title: Biomarkers and chronic GVHD

Keywords: Graft-versus-host disease, biomarkers

Word Count- Abstract: 301; **Text:** 3220

ABSTRACT

Biology-based markers to confirm the diagnosis of chronic GVHD (cGVHD) or monitor progression of this frequent long-term complication of allogeneic hematopoietic cell transplantation (HCT) are critically needed to facilitate evaluation of new therapies. Biomarkers have been defined as any characteristic that is objectively measured and evaluated as an indicator of a normal biological or pathogenic process, a pharmacologic response to a therapeutic intervention with "objectively" further described as "reliably and accurately" [1, 2]. Potential applications of biomarkers in cGVHD clinical trials or patient management include: a) diagnosis and assessment of cGVHD disease activity, including distinguishing irreversible damage from continued disease activity; b) prognostic risk to develop cGVHD; and c) prediction of response of cGVHD response to therapy and d) as a surrogate endpoint intended to substitute for, or complement, a clinical endpoint..

Approaches to identification of biomarkers include hypothesis-driven testing in pre-clinical or clinical settings, and high-throughput discovery-based methods to analyze large numbers of samples in clinical trials. Sample collection for cGVHD biomarkers studies should be well-documented with identical or comparable protocols for sample acquisition, processing, preservation and testing, at intervals that are both calendar and event driven. The consistent therapeutic treatment of subjects and standardized documentation needed to support biomarker studies are most likely to be provided in prospective clinical trials. For such multi-center clinical trials, use of central processing facilities and testing is recommended.

To date, no biomarkers for cGVHD have been qualified and utilized in the clinic. However, since our previous cGVHD Biomarkers Working Group report in 2005 [3], and increasing number

of novel studies of the biology of cGVHD have suggested candidate biomarkers for further investigation. The Biomarker Working Group of the NIH Consensus Development Project on Criteria for Clinical Trials in cGVHD is therefore optimistic that biomarkers clinically relevant in cGVHD are closer to being realized.

BACKGROUND

Chronic GVHD (cGVHD) is now one of the most important long term complications of allogeneic blood and marrow transplantation resulting in significant morbidity and mortality [4]. Unlike acute GVHD, cGVHD is insidious in its onset and the diagnosis can be difficult. Moreover, cGVHD is a multifactorial disease that can affect almost all organs and tissues in the body. Thus the identification and verification of biomarkers for cGVHD is more difficult as compared to acute GVHD (aGVHD). The new scoring system proposed by NIH Consensus Development Project on Criteria for Clinical Trials in cGVHD [5] has now been widely adopted by HCT centers. However, clinical characteristics are not fully informative in predicting the severity of the disease, response to therapy, or survival and are not adequate to distinguish disease activity from irreversible tissue damage during treatment [6]. As an adjunct to clinical and histological criteria, the availability of biomarkers for cGVHD could potentially improve the classification of patients into risk groups and thereby refine cGVHD diagnosis, or predict risk of developing cGVHD or response to therapy.

The pathogenesis of cGVHD continues to be a subject of debate. Historically, cGVHD was believed to be a chronic continuation of the same effector mechanisms that are causative of acute GVHD, i.e. donor T lymphocytes having specificity for recipient restricted histocompatibility antigens and their associated cytokines. However, differences in the clinical presentations of acute and chronic GVHD have suggested the effector mechanisms might differ. Further, therapies that are clinically effective in treating aGVHD are much less effective for cGVHD. Some investigators have hypothesized that chronic GVHD might be an autoimmune disease rather than an alloimmune disease, because of the clinical similarities of cGVHD to some human autoimmune diseases. Details of the possible pathophysiologic mechanisms are detailed in our companion manuscript [7].

Resources for biomarker development may include observational correlations with certain phenotypes. While some correlations might not reflect the underlying pathology, others will suggest new pathophysiologic pathways and potential therapeutic targets. Alternatively potential biomarkers may be identified via biological hypotheses generated from pre-clinical or theoretical models of cGVHD pathophysiology, as described in our companion manuscript [7]. However, a biomarker does not necessarily directly represent the biology of cGVHD.

Purpose of this document

To facilitate the identification, verification, qualification and application of cGVHD biomarkers, in this document we recommend: a) standard definitions for biomarkers and their applications; b) characteristics that should be included in clinical study endpoints when investigating cGVHD biomarkers; c) confounding factors that must be considered when measuring different types of cGVHD biomarkers; d) critical concepts and a recommended decision process specific to the selection and development of biomarkers in cGVHD; and e) considerations for a repository with the minimal essential clinical data to be provided with each sample.

Summary of 2014 changes

This document and our companion manuscript [7] replace the 2006 report [3] of the Biomarker Working Group of the NIH Consensus Development Project on Criteria for Clinical Trials in cGVHD.

SUMMARY OF RECOMMENDATIONS

The Biomarker Working Group makes the following recommendations:

1. Biomarker(s) of cGVHD should meet all of the following requirements:

a. Diagnostic, prognostic, or predictive of the potential to respond to a treatment, or showing response to a treatment that will lead to an important clinical outcome.
b. Methodology for measuring the biomarker is accepted as rigorous as outlined in this manuscript.
c. Evaluation in two or more independent cohorts has been completed, each having sufficient power for statistical significance.
d. Confirmation by two or more research groups has been completed.

We specify that studies be conducted in at least two cohorts and by at least two distinct research groups as the minimal acceptable criterion to confirm the validity of a particular biomarker. Moreover, analysis of biomarkers must be evaluated in similar cGVHD presentations using identical laboratory assays. This is because the observation of a significant association in a single data set does not ensure that the findings can be generalized to other data sets or that the association is specific for the investigated condition. Most biomarkers with promising results in a first data set do not hold their promise in independent data sets. Thus, if a biomarker is confirmed to have strength of association for the investigated condition in at least two completely independent cohorts and by at least two research groups, such a biomarker would have potential for use in cGVHD clinical trials or patient management. To best achieve these goals, a multi-disciplinary, coordinated approach to the identification, verification, qualification, and application of biomarkers should be implemented, particularly within the context of clinical therapeutic trials.

2. Both hypothesis-driven and discovery-based approaches for identification of cGVHD biomarkers are needed.
3. To leverage resources, both cGVHD observational prospective studies such as the one proposed below and cGVHD clinical therapeutic trials should include correlative biological studies focused on the identification, verification, qualification and application of biomarkers

whenever possible. The main advantage of observational prospective studies is their ability to capture the natural history of the cGVHD population, with the opportunity to identify and evaluate biomarkers not related to treatment response. In the proposal below, heterogeneity problems and center effects are also addressed (through multi-institutional accrual). Randomized controlled clinical trials have the advantage that baseline equality is assured in the compared groups by means of random allocation of subjects to the treatment or comparison group, thus making potential sources of bias easier to anticipate and control for. This is essential for rigorous evaluation of biomarkers that are intended to predict or show a response to treatment.

4. Samples from well-documented cases with and without cGVHD and with or without prior aGVHD should be stored using standardized protocols as proposed in this consensus paper, in order to create a resource for future biomarker studies. Samples to be collected on cases and controls and minimal essential clinical data to be provided with each sample are detailed below.

DEFINITIONS OF BIOMARKERS AND THEIR APPLICATIONS

A biomarker has been defined by the NIH Biomarker Working Group as a characteristic that is <u>objectively measured</u> as an indicator of normal biological processes or pathogenic processes, or biological responses to a therapeutic intervention [1]. The Institute of Medicine has further defined "objectively" to mean "reliably and accurately" [2]. Furthermore, for the purposes of this document, certain evaluations that are routinely performed to determine the diagnosis of cGVHD or to assess the severity of the disease are not acceptable as biomarkers. Examples of such evaluations include pulmonary function testing, and radiologic assessment including computed tomography scans.

Applications of cGVHD biomarkers critical to clinical care and research studies are summarized below and in **Table 1**.

1. *Diagnose chronic GVHD*. For example, a biomarker could be used together with clinical criteria to determine eligibility for a clinical trial or clarify differential diagnosis (e.g. infection, drug reaction, other inflammatory disease vs. cGVHD). *Distinguish cumulative damage or irreversible tissue damage from current cGVHD activity*. Many of the organ systems involved in cGVHD undergo cumulative tissue damage and grading scales do not distinguish well between the extent of current areas of inflammatory activity (e.g. infiltrates of lymphocytes into tissue) and cumulative damage (sclerotic scarring, loss of lacrimal or salivary function due to loss of secretory acini).
2. *Prognostic risk to develop cGVHD*. For example, gene polymorphisms in either the donor or recipient may be associated with risk of cGVHD. Identifying prognostic markers for the most severe forms of cGVHD, prior to onset of cGVHD, is of particular interest. *Assess risk for progression or establish staging of cGVHD*. For example, a biomarker could be used to determine the risk category for cGVHD to lead to severe sequelae or to guide decisions about the need for treatment.
3. *Predict potential for response to therapy*. For example, a biomarker may distinguish between different pathophysiologic processes that can drive cGVHD in a particular patient, and aid determining which treatment(s) are more likely to provide benefit for that specific patient. This would serve as a guide to treatment selection prior to administering treatments.
4. *Demonstrate a response to treatment, particularly a response related to a long term outcome such as nonrelapse mortality*. For example, a biomarker could be developed to monitor for therapeutic response. This type of biomarker could also help guide choices of

treatment in an empiric manner, by revealing that a treatment has not resulted in an adequate response, and a change in treatment is warranted.

Biomarkers that could be used to predict response to treatment, measure disease activity or distinguish reversible disease activity from irreversible damage would have very high clinical utility, since currently available clinical tools are not adequate for these purposes. In addition, biomarkers that prognostic risk to develop severe cGVHD would also have high utility and could be used in preemptive trials.

RISK FACTORS AND COVARIATES TO CONSIDER

While a biomarker may provide valuable assessment of cGVHD, other contributing, confounding conditions must also be considered. Some potential confounding factors are defined below and summarized in **Table 2.**

Factors affecting biomarkers directly and independent of onset of cGVHD

The conditions of a) immune reconstitution, b) concomitant aGVHD, c) the type and intensity of current immunosuppressive therapy, d) presence of infections, and e) sample processing and storage, will all impact directly on cGVHD biomarker's expression or measurement and interpretation. In particular, the analysis of immune-related biomarker data must also account for time from transplant, since immune reconstitution occurs gradually. Organ involvement by cGVHD (type of tissues involved and NIH score) and the clinical presentation at onset will provide the cGVHD diagnostic phenotype, and this may directly affect the data on biomarkers. Because of the heterogeneity of the clinical syndromes of cGVHD, it is unlikely that biomarkers applicable to all forms or presentations of this disorder will be identified.

Covariates and potential confounding factors

A variety of confounding factors may limit the interpretation of results of cGVHD biomarker studies. Each of the following confounding factors may limit the scope and application of a particular biomarker: a) recipient characteristics such as age; b) donor characteristics including treatment of the donor with G-CSF or other agent (i.e., plerixafor); and if the graft is manipulated, for example, a T cell depleted haplo-identical transplant with or without cyclophosphamide post-transplant; c) donor source and the type of graft (peripheral blood, bone marrow, or umbilical cord blood); and d) recipient preparative conditioning regimen.

CRITICAL CONCEPTS AND DECISION PROCESS AS APPLIED TO BIOMARKERS IN cGVHD

So far, most potential cGVHD biomarkers have been identified based on evaluation at a single center or single laboratory, and have not been verified and qualified; only a few have included patients derived from multiple centers or independent cohorst of patients. Thus, we propose a four-part framework for the development of cGVHD biomarkers (**Figure 1**). These recommendations are based on guidance for biomarker development from the Institute of Medicine [2] and the US Food and Drug Administration (FDA) in the Center for Drug Evaluation and Research (CDER) [8]. Careful application of these recommendations and avoidance of some of the cancer community's previous mistakes such as 1) strong reliance on convenience samples rather than a prospectively defined population in which specimens are to be collected, 2) absence of a confirmation set that is totally independent of the hypothesis-generating data, 3) lack of understanding of what it might take to develop an algorithm, or rule that combines multiple biomarkers, and 4) lack of understanding that biomarkers should provide information additive over the commonly available information on the patient, will allow our HCT community to move forward and translate biomarkers into the clinic [9,

10]. We have used the new recommended terminology [1, 2] to replace the term "validation", which is thought to be confusing. In **Figure 1** we propose an algorithm and the steps are explained below.

Phase 1: Identification. The initial phase is the identification of candidate biomarkers in a small experiment of well-matched cohorts from cases / controls selected from the populations in which it is hoped to be used. At this initial step it is important to define the clinical context of use and which Clinician Reported Outcome (CRO) or Patient Reported Outcome (PRO) data will be captured to assess a clinical endpoint, for example, nonrelapse mortality (NRM) or relapse mortality (RM) or more robust scales (such as the NIH cGVHD 0-3 organ score and the Lee cGVHD symptom scale) (we will have the new references). The most appropriate controls for the study in question should be defined at this point. It should not be assumed that the same controls will be useful for a different clinical context of use. Factors that should be considered in the choice of controls are discussed in **Table 2**.

Phase 2: Verification. This step confirms the analytical validity of a test. This includes its reproducibility, and accuracy (% coefficient of variation, precision). Test practicality should also be considered: Is the potential sample to be measured easy to obtain, is the test reproducible, and is it cost-effective? Of note, prior to the qualification step, parameters such as cutoff values and sample collection procedures are locked down (finalized) and cannot be changed without re-verification of the test under the revised conditions.

Phase 3: Qualification. This step assesses the robustness of the test in all samples from the intended use population for a certain CRO/PRO (i.e. correctness). Statistical considerations for this step are shown in **Table 3**. Other statistical analyses that have been proposed to estimate the performance of biomarkers are reviewed by Pepe et al. [11]. The Qualification Cohort for Phase 3 should be entirely

distinct and separate — different site(s), and different demographics — from the Identification Cohort previously studied in Phase 1.

Phase 4: Application. In this final step, the test is used in the clinic (e.g. in the context of all comers suspected of cGVHD) or in a prospective randomized clinical trial, and with potential to impact CRO/PRO. If the biomarker was 'proven' to have some value in the Qualification step, experience with application of the biomarker could test for example: (i) practicality of use in a consortium study, (ii) replacement of the clinical scoring system or invasive biopsies by a simple blood test, (iii) early surrogate indicator of treatment when testing a new drug as compared to the standard of care. This requires Institutional Review Board (IRB) approval. Investigational Device Exemption (IDE) or Investigational New Drug (IND) approval may also be needed, depending on whether patients in the study are managed based on the outcome of the test. This Prospective Cohort should include a larger population than evaluated for the previous cohorts. This step would best be performed as a component of a Blood and Marrow Transplant Clinical Trial Network (BMT CTN) protocol.

SAMPLE REPOSITORY FOR INVESTIGATION OF CGVHD BIOMARKERS

The largest barrier to new cGVHD biomarker development is the lack of good quality biological samples linked to granular, detailed clinical data. Well-conducted large multicenter observation or interventional clinical trials represent excellent formats to provide the consistency of standardized documentation needed to support qualification studies correlating biomarkers with clinical endpoints of interest. However, single institution or observational studies with limited institutional participation in which standardized diagnostic criteria are employed may be sufficient for initial identification studies. Thus, we propose the following:

1) Prospective multicenter studies with collection and banking of samples with a link to patient data regarding cGVHD, in a manner that complies with regulations for disclosure of protected

health information. Assessors will require adequate training to correctly collect the clinical data. The cGVHD focused clinical variables of interest are outlined in the other companion NIH Consensus Development Project on Criteria for Clinical Trials in cGVHD working group reports [5, 6]. **Table 4** shows the minimal recommended data elements that should be linked with each stored sample.

2) Sample acquisition protocols which incorporate both calendar driven timepoints and event driven sample collection. Examples of event-driven sample collection include when the patient is first diagnosed with cGVHD (or within 2 months) or before start of systemic treatment or at the time of change in treatment. Since the immune environment changes with post-transplant immune reconstitution, time-matched samples should also be obtained from patients who do not have cGVHD. In the absence of cGVHD, we recommend samples should be obtained at 3, 6, 9, 12, and 18-24 months after transplantation and then yearly, if possible up to 6-8 years. This schedule will adequately sample during the period of greatest risk for development of cGVHD, and also allow long-term studies after cGVHD treatment.
3) A centralized repository at the National Marrow Donor Program (NMDP) or possible a virtual repository with multiple sites collecting in a standardized manner. Collection of samples for the investigation of cGVHD and data acquisition in the context of other ongoing trials that collect comprehensive data before day +100, specifically BMT CTN 1201, would leverage resources most efficiently. Another advantage of this approach is the central standardization of processing and assays. Several recent publications indicate the value of standardized procedures for sample acquisition, handling and storage, and core laboratories for testing in multi-center clinical trials to ensure comparability of results [12-15].
4) The BMT CTN 1201 does not include event-driven samples or fluids/biopsies of cGVHD target organs such as bronchial lavage fluid (BAL), skin, intestine, liver, and mucosa; if

resources to collect these are not available, dedicated large cores/clinical centers should continue to collect and bank these other quality samples locally.

5) In addition, for hypothesis driven or discovery based studies, the acquisition of samples as described above should be continued.
6) Finally, subject permission to use banked samples in future research investigations and to exchange materials with other institutions should be embedded in the approved consent documents. This will allow such studies to be conducted in the future without the need for explicit re-consenting of patients.

CANDIDATE BIOMARKERS IN cGVHD

A limited number of potential cGVHD biomarkers have been evaluated in both hypothesis-driven and discovery-based testing for specific clinical applications. The data have come primarily from single centers or from a number of collaborating centers; in most cases, the findings have not been assessed as part of large multicenter trials. Despite extensive prior investigation, few potential biomarkers have been rigorously tested and replicated in independent large cohorts of patients as recommended by this consensus [16]. In **Table 5** we present published candidate cGVHD markers, organized by application (diagnostic, prognostic/risk stratification, predictive), and then in ascending strata based on the strength of the published evidence. This table illustrates how very few markers have been identified from studies incorporating discovery and independent qualification, and the relative lack of studies in the realm of cGVHD therapeutic response. Among potential biomarker applications, we emphasize prognostic/risk stratification and predictive biomarkers as major priorities for future investigation. As a reminder, biomarkers are observational correlations and might not necessarily

reflect the underlying cGVHD pathology. However, they often do, and the biology of the markers listed in **Table 5** is further described in our companion manuscript [7].

In conclusion, although progress has been made, much work will be required to verify and qualify the candidate biomarkers identified in previous studies, and to implement high-throughput methods with appropriately collected specimens for future discovery-based approaches. Close coordination between multi-specialty clinical and laboratory-based groups, as well as funding agencies and partners, will be needed to pursue such studies successfully. We are confident that identification, verification and qualification of biomarkers will greatly assist the evaluation of new approaches for treating cGVHD.

ACKNOWLEDGMENTS

The opinions expressed are those of the authors and do not represent the position of the National Cancer Institute, the National Heart, Lung and Blood Institute, the National Institute of Allergy and Infectious Diseases, the National Institutes of Health, the Food and Drug Administration, or the United States Government.

REFERENCES

[1] Biomarkers Definitions Working G. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69:89-95.
[2] Perspectives on Biomarker and Surrogate Endpoint Evaluation:Discussion Forum Summary: The National Academies Press; 2011.
[3] 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 Transpl. 2006;12:126-37.
[4] Wingard JR, Majhail NS, Brazauskas R, Wang Z, Sobocinski KA, Jacobsohn D, et al. Long-term survival and late deaths after allogeneic hematopoietic cell transplantation. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011;29:2230-9.
[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. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2005;11:945-56.
[6] 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 transplantation. 2014;49:532-8.
[7] Cooke K, Miklos DB, Sarantopoulos S, Fowler D, Gress R, Van den Brink MR, et al. Biology of chronic GVHD to test novel biomarkers and therapies. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2015.
[8] U.S. Food and Drug Administration, Center for Drug Evaluation and Research. Biomarker Qualification Program. Web. January 2014.
http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm284076.htm.
[9] Ransohoff DF. Rules of evidence for cancer molecular-marker discovery and validation. Nature reviews Cancer. 2004;4:309-14.
[10] Ransohoff DF, Gourlay ML. Sources of bias in specimens for research about molecular markers for cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28:698-704.
[11] Pepe MS. The Statistical Evaluation of Medical Tests for Classification and Prediction. New York: Oxford University Press; 2003.
[12] Maecker HT, McCoy JP, Jr., Amos M, Elliott J, Gaigalas A, Wang L, et al. A model for harmonizing flow cytometry in clinical trials. Nature immunology. 2010;11:975-8.
[13] Maecker HT, McCoy JP, Nussenblatt R. Standardizing immunophenotyping for the Human Immunology Project. Nature reviews Immunology. 2012;12:191-200.
[14] Biancotto A, Feng X, Langweiler M, Young NS, McCoy JP. Effect of anticoagulants on multiplexed measurement of cytokine/chemokines in healthy subjects. Cytokine. 2012;60:438-46.
[15] Asare AL, Kolchinsky SA, Gao Z, Wang R, Raddassi K, Bourcier K, et al. Differential gene expression profiles are dependent upon method of peripheral blood collection and RNA isolation. BMC genomics. 2008;9:474.
[16] Pidala J, Sarwal M, Roedder S, Lee SJ. Biologic markers of chronic GVHD. Bone marrow transplantation. 2014;49:324-31.
[17] Kitko CL, Levine JE, Storer BE, Chai X, Fox DA, Braun TM, et al. Plasma CXCL9 elevations correlate with chronic GVHD diagnosis. Blood. 2014;123:786-93.
[18] Kim DH, Lee NY, Sohn SK, Baek JH, Kim JG, Suh JS, et al. IL-10 promoter gene polymorphism associated with the occurrence of chronic GVHD and its clinical course during systemic

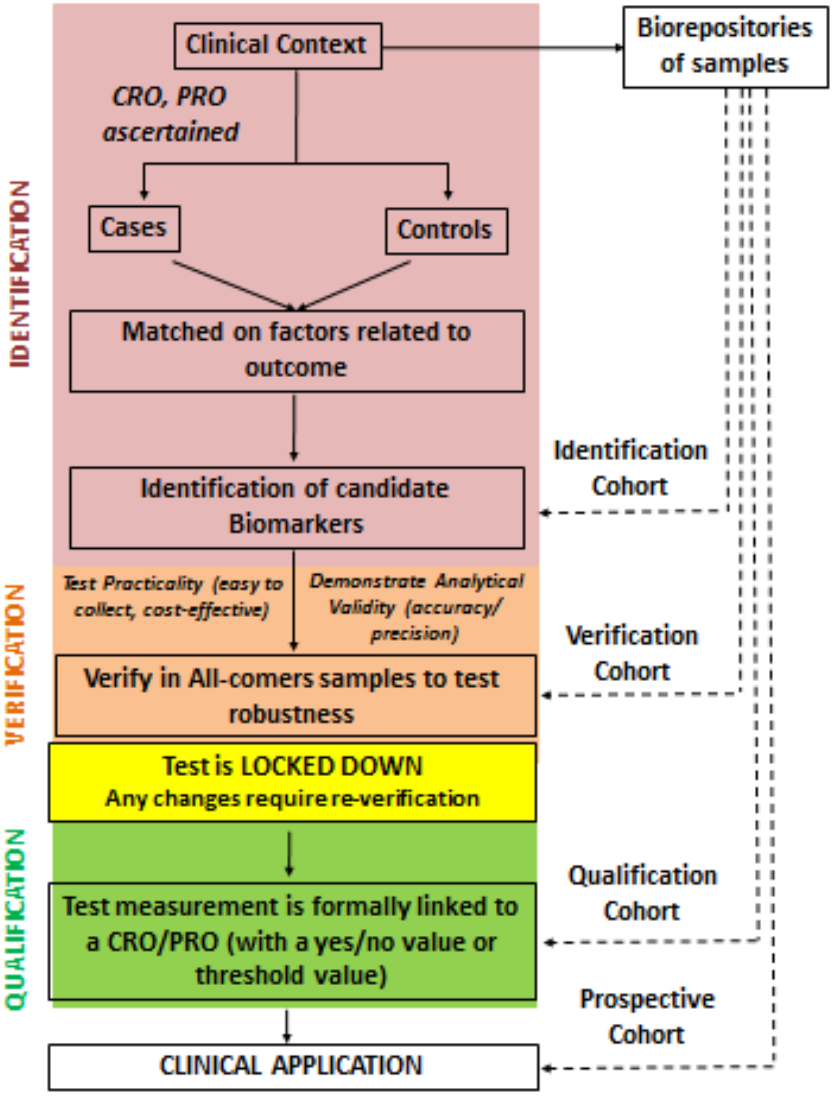
immunosuppressive treatment for chronic GVHD after allogeneic peripheral blood stem cell transplantation. Transplantation. 2005;79:1615-22.
[19] Mullighan C, Heatley S, Doherty K, Szabo F, Grigg A, Hughes T, et al. Non-HLA immunogenetic polymorphisms and the risk of complications after allogeneic hemopoietic stem-cell transplantation. Transplantation. 2004;77:587-96.
[20] Rocha V, Franco RF, Porcher R, Bittencourt H, Silva WA, Jr., Latouche A, et al. Host defense and inflammatory gene polymorphisms are associated with outcomes after HLA-identical sibling bone marrow transplantation. Blood. 2002;100:3908-18.
[21] Sivula J, Turpeinen H, Volin L, Partanen J. Association of IL-10 and IL-10Rbeta gene polymorphisms with graft-versus-host disease after haematopoietic stem cell transplantation from an HLA-identical sibling donor. BMC Immunol. 2009;10:24.
[22] Takahashi H, Furukawa T, Hashimoto S, Suzuki N, Kuroha T, Yamazaki F, et al. Contribution of TNF-alpha and IL-10 gene polymorphisms to graft-versus-host disease following allo-hematopoietic stem cell transplantation. Bone marrow transplantation. 2000;26:1317-23.
[23] Abrahamsen IW, Somme S, Heldal D, Egeland T, Kvale D, Tjonnfjord GE. Immune reconstitution after allogeneic stem cell transplantation: the impact of stem cell source and graft-versus-host disease. Haematologica. 2005;90:86-93.
[24] D'Orsogna LJ, Wright MP, Krueger RG, McKinnon EJ, Buffery SI, Witt CS, et al. Allogeneic hematopoietic stem cell transplantation recipients have defects of both switched and igm memory B cells. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2009;15:795-803.
[25] Greinix HT, Pohlreich D, Kouba M, Kormoczi U, Lohmann I, Feldmann K, et al. Elevated numbers of immature/transitional CD21- B lymphocytes and deficiency of memory CD27+ B cells identify patients with active chronic graft-versus-host disease. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2008;14:208-19.
[26] Kuzmina Z, Greinix HT, Weigl R, Kormoczi U, Rottal A, Frantal S, et al. Significant differences in B-cell subpopulations characterize patients with chronic graft-versus-host disease-associated dysgammaglobulinemia. Blood. 2011;117:2265-74.
[27] Sarantopoulos S, Stevenson KE, Kim HT, Cutler CS, Bhuiya NS, Schowalter M, et al. Altered B-cell homeostasis and excess BAFF in human chronic graft-versus-host disease. Blood. 2009;113:3865-74.
[28] She K, Gilman AL, Aslanian S, Shimizu H, Krailo M, Chen Z, et al. Altered Toll-like receptor 9 responses in circulating B cells at the onset of extensive chronic graft-versus-host disease. Biol Blood Marrow Transpl. 2007;13:386-97.
[29] Miura Y, Thoburn CJ, Bright EC, Phelps ML, Shin T, Matsui EC, et al. Association of Foxp3 regulatory gene expression with graft-versus-host disease. Blood. 2004;104:2187-93.
[30] Zorn E, Kim HT, Lee SJ, Floyd BH, Litsa D, Arumugarajah S, et al. Reduced frequency of FOXP3+ CD4+CD25+ regulatory T cells in patients with chronic graft-versus-host disease. Blood. 2005;106:2903-11.
[31] Matsuoka K, Kim HT, McDonough S, Bascug G, Warshauer B, Koreth J, et al. Altered regulatory T cell homeostasis in patients with CD4+ lymphopenia following allogeneic hematopoietic stem cell transplantation. The Journal of clinical investigation. 2010;120:1479-93.
[32] Rozmus J, Schultz KR, Wynne K, Kariminia A, Satyanarayana P, Krailo M, et al. Early and late extensive chronic graft-versus-host disease in children is characterized by different Th1/Th2 cytokine profiles: findings of the Children's Oncology Group Study ASCT0031. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2011;17:1804-13.

[33] Tanaka J, Imamura M, Kasai M, Hashino S, Kobayashi S, Noto S, et al. Th2 cytokines (IL-4, IL-10 and IL-13) and IL-12 mRNA expression by concanavalin A-stimulated peripheral blood mononuclear cells during chronic graft-versus-host disease. European journal of haematology. 1996;57:111-3.
[34] Dander E, Balduzzi A, Zappa G, Lucchini G, Perseghin P, Andre V, et al. Interleukin-17-producing T-helper cells as new potential player mediating graft-versus-host disease in patients undergoing allogeneic stem-cell transplantation. Transplantation. 2009;88:1261-72.
[35] Barak V, Levi-Schaffer F, Nisman B, Nagler A. Cytokine dysregulation in chronic graft versus host disease. Leukemia Lymphoma. 1995;17:169-73.
[36] Li Q, Zhai Z, Xu X, Shen Y, Zhang A, Sun Z, et al. Decrease of CD4(+)CD25(+) regulatory T cells and TGF-beta at early immune reconstitution is associated to the onset and severity of graft-versus-host disease following allogeneic haematogenesis stem cell transplantation. Leukemia research. 2010;34:1158-68.
[37] Skert C, Damiani D, Michelutti A, Patriarca F, Arpinati M, Fil C, et al. Kinetics of Th1/Th2 cytokines and lymphocyte subsets to predict chronic GVHD after allo-SCT: results of a prospective study. Bone marrow transplantation. 2009;44:729-37.
[38] Ritchie D, Seconi J, Wood C, Walton J, Watt V. Prospective monitoring of tumor necrosis factor alpha and interferon gamma to predict the onset of acute and chronic graft-versus-host disease after allogeneic stem cell transplantation. Biology of Blood and Marrow Transplantation. 2005;11:706-12.
[39] Cavet J, Dickinson AM, Norden J, Taylor PR, Jackson GH, Middleton PG. Interferon-gamma and interleukin-6 gene polymorphisms associate with graft-versus-host disease in HLA-matched sibling bone marrow transplantation. Blood. 2001;98:1594-600.
[40] Cullup H, Dickinson AM, Cavet J, Jackson GH, Middleton PG. Polymorphisms of interleukin-1alpha constitute independent risk factors for chronic graft-versus-host disease after allogeneic bone marrow transplantation. British journal of haematology. 2003;122:778-87.
[41] Visentainer JEL, Lieber SR, Persoli LBL, Vigorito AC, Aranha FJP, de Brito Eid KA, et al. Serum cytokine levels and acute graft-versus-host disease after HLA-identical hematopoietic stem cell transplantation. Experimental hematology. 2003;31:1044-50.
[42] Bertinetto FE, Dall'Omo AM, Mazzola GA, Rendine S, Berrino M, Bertola L, et al. Role of non-HLA genetic polymorphisms in graft-versus-host disease after haematopoietic stem cell transplantation. Int J Immunogenet. 2006;33:375-84.
[43] Viel DO, Tsuneto LT, Sossai CR, Lieber SR, Marques SB, Vigorito AC, et al. IL2 and TNFA gene polymorphisms and the risk of graft-versus-host disease after allogeneic haematopoietic stem cell transplantation. Scandinavian journal of immunology. 2007;66:703-10.
[44] Zorn E, Mohseni M, Kim H, Porcheray F, Lynch A, Bellucci R, et al. Combined CD4+ donor lymphocyte infusion and low-dose recombinant IL-2 expand FOXP3+ regulatory T cells following allogeneic hematopoietic stem cell transplantation. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2009;15:382-8.
[45] Koreth J, Matsuoka K, Kim HT, McDonough SM, Bindra B, Alyea EP, 3rd, et al. Interleukin-2 and regulatory T cells in graft-versus-host disease. The New England journal of medicine. 2011;365:2055-66.
[46] Matsuoka K, Koreth J, Kim HT, Bascug G, McDonough S, Kawano Y, et al. Low-dose interleukin-2 therapy restores regulatory T cell homeostasis in patients with chronic graft-versus-host disease. Science translational medicine. 2013;5:179ra43.
[47] Fujii H, Cuvelier G, She K, Aslanian S, Shimizu H, Kariminia A, et al. Biomarkers in newly diagnosed pediatric-extensive chronic graft-versus-host disease: a report from the Children's Oncology Group. Blood. 2008;111:3276-85.

[48] Kobayashi S, Imamura M, Hashino S, Tanaka J, Asaka M. Clinical relevance of serum soluble interleukin-2 receptor levels in acute and chronic graft-versus-host disease. Leukemia & lymphoma. 1997;28:159-69.
[49] D'Asaro M, Dieli F, Caccamo N, Musso M, Porretto F, Salerno A. Increase of CCR7- CD45RA+ CD8 T cells (T(EMRA)) in chronic graft-versus-host disease. Leukemia. 2006;20:545-7.
[50] Yamashita K, Choi U, Woltz PC, Foster SF, Sneller MC, Hakim FT, et al. Severe chronic graft-versus-host disease is characterized by a preponderance of CD4(+) effector memory cells relative to central memory cells. Blood. 2004;103:3986-8.
[51] Arpinati M, Chirumbolo G, Marzocchi G, Baccarani M, Rondelli D. Increased donor CD86+CD14+ cells in the bone marrow and peripheral blood of patients with chronic graft-versus-host disease. Transplantation. 2008;85:1826-32.
[52] Kohrt HE, Tian L, Li L, Alizadeh AA, Hsieh S, Tibshirani RJ, et al. Identification of gene microarray expression profiles in patients with chronic graft-versus-host disease following allogeneic hematopoietic cell transplantation. Clin Immunol. 2013;148:124-35.
[53] Westekemper H, Meller S, Citak S, Schulte C, Steuhl KP, Homey B, et al. Differential chemokine expression in chronic GVHD of the conjunctiva. Bone marrow transplantation. 2010;45:1340-6.
[54] McGuirk J, Hao G, Hou W, Abhyankar S, Williams C, Yan W, et al. Serum proteomic profiling and haptoglobin polymorphisms in patients with GVHD after allogeneic hematopoietic cell transplantation. Journal of hematology & oncology. 2009;2:17.
[55] Liem L, van Houwelingen H, Goulmy E. Serum Cytokine Levels after HLA-identical Bone Marrow Transplantation. Transplantation. 1998;66:863-71.
[56] Bassim CW, Ambatipudi KS, Mays JW, Edwards DA, Swatkoski S, Fassil H, et al. Quantitative salivary proteomic differences in oral chronic graft-versus-host disease. Journal of clinical immunology. 2012;32:1390-9.
[57] Devic I, Shi M, Schubert MM, Lloid M, Izutsu KT, Pan C, et al. Proteomic Analysis of Saliva from Patients with Oral Chronic Graft-Versus-Host Disease. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2014.
[58] Wechalekar A, Cranfield T, Sinclair D, Ganzckowski M. Occurrence of autoantibodies in chronic graft vs. host disease after allogeneic stem cell transplantation. Clinical and laboratory haematology. 2005;27:247-9.
[59] Svegliati S, Olivieri A, Campelli N, Luchetti M, Poloni A, Trappolini S, et al. Stimulatory autoantibodies to PDGF receptor in patients with extensive chronic graft-versus-host disease. Blood. 2007;110:237-41.
[60] Miklos DB, Kim HT, Miller KH, Guo L, Zorn E, Lee SJ, et al. Antibody responses to H-Y minor histocompatibility antigens correlate with chronic graft-versus-host disease and disease remission. Blood. 2005;105:2973-8.
[61] Rajasekar R, Mathews V, Lakshmi KM, Sellathamby S, George B, Viswabandya A, et al. Plasmacytoid dendritic cell count on day 28 in HLA-matched related allogeneic peripheral blood stem cell transplant predicts the incidence of acute and chronic GVHD. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2008;14:344-50.
[62] Larghero J, Rocha V, Porcher R, Filion A, Ternaux B, Lacassagne MN, et al. Association of bone marrow natural killer cell dose with neutrophil recovery and chronic graft-versus-host disease after HLA identical sibling bone marrow transplants. British journal of haematology. 2007;138:101-9.
[63] Yamasaki S, Henzan H, Ohno Y, Yamanaka T, Iino T, Itou Y, et al. Influence of transplanted dose of CD56+ cells on development of graft-versus-host disease in patients receiving G-CSF-mobilized peripheral blood progenitor cells from HLA-identical sibling donors. Bone marrow transplantation. 2003;32:505-10.

[64] Sarantopoulos S, Stevenson KE, Kim HT, Bhuiya NS, Cutler CS, Soiffer RJ, et al. High Levels of B-Cell Activating Factor in Patients with Active Chronic Graft-Versus-Host Disease. Clinical cancer research : an official journal of the American Association for Cancer Research. 2007;13:6107-14.
[65] Shimura K, Ashihara E, Shimazaki C, Matsunaga S, Taniguchi K, Uchiyama H, et al. Circulating endothelial progenitor cells decreased in patients with sclerodermatous chronic graft-versus-host disease. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2008;14:426-37.
[66] Pratt LM, Liu Y, Ugarte-Torres A, Hoegh-Petersen M, Podgorny PJ, Lyon AW, et al. IL15 levels on day 7 after hematopoietic cell transplantation predict chronic GVHD. Bone marrow transplantation. 2013;48:722-8.
[67] Clark WB, Brown-Gentry KD, Crawford DC, Fan KH, Snavely J, Chen H, et al. Genetic variation in recipient B-cell activating factor modulates phenotype of GVHD. Blood. 2011;118:1140-4.
[68] Broen K, van der Waart AB, Greupink-Draaisma A, Metzig J, Feuth T, Schaap NP, et al. Polymorphisms in CCR6 are associated with chronic graft-versus-host disease and invasive fungal disease in matched-related hematopoietic stem cell transplantation. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2011;17:1443-9.
[69] Inamoto Y, Murata M, Katsumi A, Kuwatsuka Y, Tsujimura A, Ishikawa Y, et al. Donor single nucleotide polymorphism in the CCR9 gene affects the incidence of skin GVHD. Bone marrow transplantation. 2010;45:363-9.
[70] Shimada M, Onizuka M, Machida S, Suzuki R, Kojima M, Miyamura K, et al. Association of autoimmune disease-related gene polymorphisms with chronic graft-versus-host disease. British journal of haematology. 2007;139:458-63.
[71] Ostrovsky O, Shimoni A, Rand A, Vlodavsky I, Nagler A. Genetic variations in the heparanase gene (HPSE) associate with increased risk of GVHD following allogeneic stem cell transplantation: effect of discrepancy between recipients and donors. Blood. 2010;115:2319-28.
[72] Kornblit B, Masmas T, Petersen SL, Madsen HO, Heilmann C, Schejbel L, et al. Association of HMGB1 polymorphisms with outcome after allogeneic hematopoietic cell transplantation. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2010;16:239-52.
[73] Bogunia-Kubik K, Mlynarczewska A, Wysoczanska B, Lange A. Recipient interferon-gamma 3/3 genotype contributes to the development of chronic graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. Haematologica. 2005;90:425-6.
[74] Ambruzova Z, Mrazek F, Raida L, Stahelova A, Faber E, Indrak K, et al. Possible impact of MADCAM1 gene single nucleotide polymorphisms to the outcome of allogeneic hematopoietic stem cell transplantation. Human immunology. 2009;70:457-60.
[75] Boukouaci W, Busson M, Peffault de Latour R, Rocha V, Suberbielle C, Bengoufa D, et al. MICA-129 genotype, soluble MICA, and anti-MICA antibodies as biomarkers of chronic graft-versus-host disease. Blood. 2009;114:5216-24.
[76] Arora M, Lindgren B, Basu S, Nagaraj S, Gross M, Weisdorf D, et al. Polymorphisms in the base excision repair pathway and graft-versus-host disease. Leukemia. 2010;24:1470-5.
[77] Olivieri A, Cimminiello M, Corradini P, Mordini N, Fedele R, Selleri C, et al. Long-term outcome and prospective validation of NIH response criteria in 39 patients receiving imatinib for steroid-refractory chronic GVHD. Blood. 2013;122:4111-8.

Figure 1. Example of a biomarker development project



CRO: Clinician Reported Outcome, PRO: Patient Reported Outcome

Table 1: Definitions of biomarkers

Types of biomarkers	Definition	Uses	Minimum matching criteria for control samples*
Diagnostic	An assay that identifies patients at the onset of clinical disease - Different forms of cGVHD may have different markers - Different tissues may result in different markers	To help in rapid diagnosis and direct initiation or alterations in therapy	a. Time from transplant b. Absence of relapse c. Absence or presence of current or recent aGVHD d. Absence or presence of active infection e. Absence of recent B cell depletion after BMT f. Manipulation or treatment of the donor product (i.e., T cell depletion, G-CSF)
Prognostic	An assay that categorizes patients by degree of risk for disease occurrence or progression	A prognostic biomarker provides information about the anticipated natural history of the disorder in that particular patient	b. Time from transplant # Prior aGVHD c. T or B cell depletion during conditioning
Predictive	An assay that categorizes patients by their likelihood of response to a particular treatment when measured prior to the treatment	A predictive biomarker provides information about whether a given patient is likely to respond to a treatment intervention in a particular way	b. Time from transplant Current immune therapy (e.g. glucocorticosteroids, calcineurin inhibitors)
Response to treatment	An assay measured after initiation of therapy that is intended to substitute for a clinical efficacy endpoint (note: a pre therapy sample for comparison is required)	A validated response marker that can be utilized in place of an accepted clinical response endpoint (see the NIH cGVHD response criteria paper)	a. Time from transplant b. Absence of relapse c. Absence or presence of current or recent aGVHD d. Absence or presence of active infection e. Absence of recent B cell depletion after BMT f. Manipulation or treatment of the donor product (i.e., T cell depletion, G-CSF)

* Factors listed in Table 2 should be considered in interpretation
# First 2 years: within 2 months, After 2 years: within 1 year. Appropriate time interval could be different in adult versus pediatric patients due to a faster immune reconstitution in pediatric patients

Table 2: Factors that must be considered for cGVHD studies	
Factor	**Impact on cGVHD biomarker**
Factors affecting biomarkers directly	
Immune reconstitution post BMT and time from transplant	• Some marker may vary with immune reconstitution post-HCT thus time and age matched (pediatric vs adult) controls are required.
Concomitant aGVHD	• Concurrent presence may overlap with classic cGVHD manifestations • Markers may represent late aGVHD manifestations
Current immune suppression	• Many immunosuppressive treatment particularly steroids may impact concentration of biomarkers (e.g. sBAFF)
Current infection	• Active infections may change cytokine milieu and markers. • CMV reactivation, pulmonary infections.
Sample processing and storage	• Some B cell populations are lost when processed with Ficoll • Choice of serum or heparin, EDTA or citrate plasma alters analytes • Time after blood draw reduces some analytes • Collection of the samples may be specific for the type of assay and the type of tissue collected (i.e., serum, urine, saliva) These factors apply both during the identification and confirmation of a biomarker as well as during its subsequent application.
Covariates and potential confounding factors	
Recipient characteristics	• Younger age associated with lower incidence of cGVHD • Non malignant diagnoses may impact on the incidence and type of cGVHD (particularly non-malignant disorders with marrow failure and chromosomal instability appear to have a higher rate of cGVHD). • Allo immunized patients may have a lower rate of engraftment result in split donor chimerism and impact the incidence of cGVHD • Non-HLA polymorphisms may impact on incidence or presentation (i.e., ABO incompatibility) • Markers may be organ specific
Donor characteristics	• Unrelated donor versus related • HLA mismatched versus HLA matched • Female donor has higher incidence of cGVHD • UCB, Peripheral blood, or marrow graft • Non-HLA polymorphisms may impact on incidence or presentation (i.e., ABO incompatibility) • Treatment of donor product (i.e., G-CSF, T cell or B cell depletion)
Preparative conditioning regimen	• MAC versus RIC • Use of T cell or B cell depletion (TCD, ATG, Campath 1H, Rituximab), all associated with a lower incidence of cGVHD • TBI associated with increased cutaneous sclerosis

Table 3: Statistical considerations
A. Analytical performance parameters • Precision (repeatability and reproducibility of an assay) • Accuracy • Assay sensitivity (limit of detection) • Assay specificity (interference, cross-reactivity) • Sample type and matrix • Sample preparation
B. Diagnostic accuracy • Sensitivity: Proportion of subjects in a sample of patients with the target condition in whom the test is positive. • Specificity: Proportion of subjects in a sample of patients without the target condition in whom the test is negative. • Receiver operator characteristic (ROC): A plot of the true-positive rate versus the false-positive rate for all possible thresholds of a biomarker. • Positive predictive value (PPV): Proportion of patients in the overall population with a positive test who have the target condition. • Negative predictive value (NPV): Proportion of patients in the overall population with a negative test who do not have the target condition.

Table 4: Minimal essential clinical/routine laboratory data to be provided with each sample	
Essential data	**Recommended data (will be marker specific)***
1) Clinical Phenotype - For diagnosis markers: NIH diagnosis and staging forms - For prediction of response: NIH response to treatment forms - This includes presence or not of concomitant features of aGVHD 2) Time after transplantation of cGVHD diagnosis or of time matched non-cGVHD patients 3) Current type of Immunosuppression (and for steroids: dose and weight of patient)	1) Prior aGVHD 2) MAC vs. RIC 3) Prior cGVHD 4) TCD vs. Not TCD 5) Prior Immunosuppressive therapies failed 6) PBSC vs. BM vs. UCB 7) Recent B cell depletion 8) Sex mismatch 9) HLA mismatch 10) Absolute T and B cells counts 11) IgG levels 12) Active uncontrolled infection (particularly CMV)

* Variables that could confound the analyses should be collected in minimal essential data

Table 5: Candidate cGVHD biomarkers$

	Diagnostic			Prognostic/risk stratification			Predictive		
Category	Cellular	Mediator	Antibodies	Cellular	Mediator	Gene polymorphism	Cellular	Mediator	Antibodies
Examined in 2 cohorts		BAFF[17] CXCL9[17] elafin[17] sCD13[17] sIL-2R[17]				IL-10[18-22]			
Cohesive findings in ≥ 2 reports	B cells[23-28] Treg[29-31]	IL-4[32, 33] IL-6[34, 35] TNFα[34, 35]		Treg[29-31, 36, 37]	TNFα [36-38]	IL-6[39-41] TNFα[19, 42, 43]	Treg[44-46]	sIL-2R[47, 48]	
Potential candidates	BAFF/B cell[27] CD3+ T cells[23] Effector memory (CD4+ and CD8+)[49, 50] Monocytes[51] Naïve CD8+ T cells[49] NK cells[23] Th17 cells[34] TLR-9 responsive B cells[28]	ADAMTS2[52] ADAMTS3[52] AREG[52] BCAT1[52] CPM[52] CXCL10[53] CXCR3[53] CXCR7[52] DAP2IP[52] haptoglobin[54] IFN-γ[32] IL-1β[35] IL-1Ra[55] IL-1R2[52] IL-2[32] IL-8[34] IL-10[55] IL-12[33] IRS2[52] SRGAP1[52] lactotransferin/lactoperoxidase (salivary) [56] IL-1Ra + CTSB (salivary) [57]	ANA[58] Anti-dsDNA[47] Anti-PDGFR[59] H-Y antibodies[60]	B cells[27] DC2[61] NK[62, 63]	BAFF[64] BAFF/B-cell ratio[64] b-FGF[65] IFN-γ[37] IL-10[37] IL-15[66] TGFβ[36] VEGF[65]	BAFF[67] CCR6[68] CCR9[69] FAS[19] FCRL3[70] Haptoglobin[54] Heparanase[71] HMGB1[72] IFN-γ[73] IL-1[40] IL-1Ra[20] MadCAM-1[74] MICA[75] PARP1[76]	Immature/memory B cell ratio[25] Th17 cells[34] TLR-9 responsive B cells[28]	BAFF[47]	Anti-PDGFR[77]

* The table presents a synthesis of published candidates chronic GVHD biomarkers organized according to application. Importantly, these are candidates based on current knowledge. Additional replication of these findings is needed, and none of the summarized candidates meet criteria for qualification.
* The table presents general biomarker candidates, and does not specifically present data on association between biomarker candidates and chronic GVHD organ involvement or severity.
* **Definitions:**
- Application: Diagnostic – distinguish chronic GVHD from non-GVHD controls; Risk stratification – determine risk for chronic GVHD development; Predictive – assess therapeutic response;
- Category: Cellular – immune cell populations; mediator – inflammatory or immune regulatory cytokines and other factors; antibodies – auto-antibodies detected in chronic GVHD; gene polymorphism – reported cytokine gene polymorphism associated with chronic GVHD.
- Strength of evidence (presented in ascending order): Potential candidates – summary of reported chronic GVHD biomarker candidates; Cohesive findings in ≥ 2 reports – Consistent findings (e.g. candidate biomarker is elevated in chronic GVHD patients vs. controls) across ≥ 2 studies, irrespective of methodology used in each report; Examined in 2 cohorts – Presented as highest ranking for quality of evidence, as the investigation included discovery and confirmation in a separate patient cohort (however, does not indicate that other investigators have replicated these findings in independent cohorts).
$Some of the markers have been evaluated in more than one independent patient cohort by either the original groups to identify the marker or by another laboratory. Moreover some markers have been identified by different methodologies when separate laboratories have evaluated the marker. We have not noted these differences in this table.