Title Page

Title:	National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: III. The 2020 Treatment of Chronic GVHD Report
Short Title:	NIH Consensus: Treatment of Chronic GVHD
Authors and Affiliations:	
Corresponding Author	
Abstract word count:	
Manuscript word count:	
Tables:	
Figures:	
Appendices:	
Financial Disclosure Statement:	

Abstract

Keywords: Chronic graft-versus-host disease, allogeneic hematopoietic cell transplantation

Introduction

Chronic graft-versus-host disease (GVHD) is the leading cause of late morbidity and non-relapse mortality after allogeneic hematopoietic cell transplantation (HCT). Treatment of chronic GVHD is often ineffective, with frequent incomplete responses and recurrences.[1, 2] Since the previous National Institutes for Health (NIH) Consensus Conference in 2014,[3] significant advances have expanded treatment options. The focus of chronic GVHD therapeutics has shifted from broad immunosuppression towards immunomodulatory agents that target pathways relevant to the pathophysiology of the disease.[4, 5] Through industry collaboration, large, multicenter phase II and phase III clinical trials have now investigated multiple oral agents, including ibrutinib (NCT02195869, NCT02959944), ruxolitinib (NCT03112603), itacitinib (NCT03584516), and belumosudil (KD025) (NCT03640481) in the treatment of upfront or steroid-refractory chronic GVHD. Based on results of a phase II clinical trial, the US Food and Drug Administration (FDA) approved ibrutinib for treatment of chronic GVHD after failure of one or more lines of systemic therapy. [6, 7] Additional drug approvals from the FDA in the treatment of chronic GVHD may be forthcoming, once results of ongoing studies become available.

These recent advances have not yet clearly indicated that outcomes of treatment for chronic GVHD have improved. The general approach to initial systemic therapy remains unchanged, as corticosteroids continue to be the standard.[8] Treatment with systemic corticosteroids has variable clinical effectiveness and can cause significant short and long-term side effects that can be as challenging as chronic GVHD itself. Furthermore, given the heterogeneity of clinical manifestations and treatment history, clinical data to guide the choice of therapy are lacking. Although novel agents have expanded treatment options,[9] the chronicity of the disease and its often

irreversible fibrotic progression remain challenging, and many patients receive multiple lines of empirically selected treatments with marginal efficacy.

New insights into the associations between clinical variables of chronic GVHD, pathophysiologic mechanisms of disease and the clinical and biologic effects of novel therapeutic agents, are required to overcome current barriers and allow for a more biological and individualized approach to the disease.

Purpose of this Document

The current report is focused primarily on setting research priorities and direction in the treatment of chronic GVHD. Future trials should be designed in a way that maximizes an understanding of both the biological and clinical effects of novel therapies. These studies should take into consideration important factors that might influence these effects, such as subject eligibility criteria, clinical trial endpoints, and safety considerations in two different clinical scenarios. Clinical trials for initial systemic treatment of chronic GVHD should be designed to determine the feasibility of minimizing or eliminating the need for concurrent corticosteroid treatment. Clinical trials for treatment-refractory disease should focus attention on situations where responses tend to be incomplete or suboptimal. Trials in both scenarios should integrate biological correlative testing. This document is not meant to be prescriptive, as it is neither a review nor a treatment guideline. Rather, we provide guidance to the overall approach of clinical trial development for treatment of chronic GVHD.

Summary of Recommendations

1. Detailed correlative scientific studies should be conducted in the context of clinical trials to evaluate associations between clinical outcomes and the biological impact of systemic therapeutics. All trials should collect samples that can be preserved for future analysis.
2. Broad panels of biological correlative studies should be evaluated initially in a small number of patients to inform the selection of a narrower panel of informative correlative measures to test in subsequent studies.
3. Biopsies of involved tissue should be obtained before starting treatment to enable evaluation of disease involvement, response to therapy, and correlation with blood biomarkers.
4. Advanced collaboration between academic medical centers, medical societies, and industry is required to support a biology-based strategic approach.
5. Clinical trials for initial systemic therapy should investigate novel agents in designs that minimize or eliminate concurrent corticosteroid administration. These trials should be conducted at experienced academic medical centers with infrastructure to support biological sample processing and analysis.
6. For patients with treatment-refractory chronic GVHD, trials targeting specific disease manifestations such as fibrosis and bronchiolitis obliterans may allow a more informative assessment of clinical and biological results than would otherwise be possible.
7. Future clinical trials should aim toward development of algorithms for a personalized approach to treatment based on the clinical phenotype and biological profile of each patient.
8. Randomized clinical trials should be conducted for agents which demonstrate disease activity in large single arm studies.

9. Master protocols should be explored to enable rapid clinical screening of new treatments in early-phase studies.

Methods

Each working group was organized to encourage global engagement in the topic (see introduction to the series). Four groups worked individually beginning in February 2020 to review the relevant literature and prepare the initial draft of the manuscript. The Steering Committee reviewed and discussed the initial draft and offered recommendations for revisions. Two iterative rounds of comments and revisions were collected before the November 2020 Consensus Conference. The manuscript was further revised for submission after additional suggestions from external reviewers, virtual Conference participants, and a 30-day public comment period.

Biological Considerations

Biological Implications

Although biomarkers could be useful in predicting chronic GVHD onset, severity, or response to therapy, such data do not necessarily provide insight into pathogenesis or reasons for response or lack of response after therapeutic interventions. The availability of biological specimens obtained before and after a study intervention in clinical trials would make it possible to draw pathophysiologic inferences from correlative changes in laboratory measures and clinical outcomes. Results of these studies could ultimately inform the development of personalized approaches in the management of chronic GVHD.

Extensive pre-clinical small animal and human ex vivo investigations have implicated certain biological pathways in the pathophysiology of chronic GVHD and helped identify therapeutic agents that may have clinical activity.[5] However, the impact of disease course and treatment on these pathways in patients with chronic GVHD are poorly understood. Thus, the association between chronic GVHD disease manifestations, such as individual organ involvement or overall disease phenotype, and biological markers of pathophysiology are not well understood and should to be elucidated in future studies. Furthermore, the biological impact of therapeutic agents and the association between biological and clinical effect is not well characterized. Preliminary investigation into the biological activity of novel agents has focused on the proposed mechanism of action. How these agents modulate other pathogenic signaling pathways in chronic GVHD is not known. Cause-and-effect testing of drug, protein/antibody, or cellular therapies can be interrogated in preclinical models chosen to best reflect clinical phenotypes and biological data.

As we better understand this clinical-biological correlation, the approach to chronic GVHD treatment could significantly change. The association of biological markers with disease manifestations could lead to algorithms that refine risk stratification, with vast implications for treatment-related outcomes and clinical trial design. Moreover, select biomarkers or panels of biomarkers could guide selection of specific therapeutic agents and be used to monitor responses with treatment.[10] For all these reasons, future studies should expand the assessment of biological correlatives in the context of clinical trial designs in order to close these knowledge gaps and redefine the paradigms of chronic GVHD treatment into a more personalized approach (**Figure 1**).

Evaluation of Biological Markers

Preclinical studies have identified numerous pathophysiological pathways in chronic GVHD. Phenotypic patterns of chronic GVHD may be classified into distinct groups based on clinical and organ system manifestations by machine learning techniques that will continue to evolve.[11] Correlation of biological markers with such phenotypic groups could enable better risk stratification and identify pathogenic mechanisms with implications for treatment.[12] With accumulated data from multiple trials over time, it may be possible to identify more precise therapeutic targeting of particular cells, pathways or tissues for a given phenotypic group. Equally important, paired analysis of biomarker changes between baseline and subsequent time points after starting treatment with a specific agent would be highly informative in documenting expected on-target and unexpected off-target pharmacodynamic effects.[13]

The data needed for such an approach can be divided into tests that are readily available at many sites and are not cost prohibitive (Tier 1) and others that require special equipment, sophisticated analytics, or complex sample processing or incur high cost (Tier 2). Biospecimens can include peripheral blood cells and aspirated marrow, bone marrow and tissue biopsies, and serum and plasma. Correlative studies testing a single agent for initial treatment are likely to be more informative than studies of treatment-refractory disease, which may be confounded by prior disease and treatment effects. By collecting samples before, during and after therapy, the feasibility and utility of an assay can be judged in a pilot study, leading to definition of a more focused panel of testing to pursue in follow-up studies.

Tier 1 studies

As upfront monotherapy drug trials would likely target different pathways, cell types and function, and tissue injury, tier 1 (and tier 2) studies should be designed to ensure that any targeted analyses are added to a widely adopted panel of baseline laboratory measures that would enable meaningful comparisons between studies. This baseline panel of laboratory measures should be adjusted as relationships with clinical phenotype and outcomes are clarified by emerging information.

Since chronic GVHD requires a cellular response, flow cytometry of peripheral blood cells could help identify cell types that may be pathogenic, regulatory, anergic or tolerant and which are associated with a distinct chronic GVHD phenotype. Multi-parameter flow cytometry should consider focusing on immunological parameters that could affect chronic GVHD pathogenesis and therapeutic targets, including T cell differentiation (i.e., Tnaive, central memory, effector memory, memory stem, regulatory T cells, Tfollicular helper cells) and function (i.e., subsets that produce inflammatory or anti-inflammatory cytokines), the capacity for cytolysis or regulation via well-defined pathways, or markers of anergy. Expression of chemokine, adressin, homing and inhibitory receptors, costimulatory ligands, and cytokine receptors could provide functional insights. Certain NK cell subsets can produce proinflammatory cytokines or contribute to tissue destruction by cytolysis, while others have regulatory functions. Therefore, collecting NK cell data could add to our understanding of the immune system alterations in chronic GVHD.

B cell developmental and maturation stages (i.e., bone marrow pro-B, large and small pre-B and immature B cell subsets, peripheral blood transitional, naïve, memory and germinal center B cells), and functions (i.e., production of auto- and alloantibodies, anti-inflammatory cytokines, activation

to support T cell responses, induction of anergy, and regulatory characteristics) could aid in understanding the extent of chronic GVHD pathogenesis and identify therapeutic targets to correct abnormalities. Short-lived plasmablasts and plasma cells and long-lived plasma cells could antibody that contributes to pathogenesis. Circulating dendritic cells and inflammatory or anti-inflammatory monocytes or macrophage lineages could be enumerated and correlated with tissue infiltrates and fibrosis.

Histopathological and immunohistochemical studies of research tissue biopsies could reveal the involved organ, location within the organ/tissue and function (cytokine production; cytolytic and antigen presentation capacity, regulatory properties) of infiltrating cells and types of injury (inflammatory; fibrotic) associated with a given clinical phenotype and laboratory measures. Serum and plasma could be analyzed for cytokines and specific chemokines that recruit pathogenic or regulatory cells into tissues. Autoimmune antibody screens can be performed using commercial Luminex or ELISA kits to help clarify the role of B cell activation across different phenotypic groups. Expression of inhibitory, costimulatory, and regulatory antigens recipient tissue cells would complete the picture.

Tier 2 Studies

These assays require storage of biospecimens in a way that does not alter analyte measurement in future assays. The Working Group recommends suitable cryopreservation of serum, plasma, cells and archival of tissue biopsy materials for future testing. Cells should be preserved in RNAlater or similar buffers. Multi-parameter CyTOF with heavy metals would allow assessment of intracellular signaling and transcription factor expression coupled with cell surface antigen

expression. Applicable high throughput, sophisticated technologies include single cell RNA-seq, CITE-seq, which measures both cell surface proteins and RNA-seq, and single cell RNA-seq with ATAC-seq, which assesses chromatin accessibility. Serum and plasma can be assessed by mass spectrometry to identify proteins or metabolites whose concentrations are altered during chronic GVHD. Tissues may be processed using multiplexed immunofluorescence to define spatial relationships between subsets of infiltrating donor-derived cells, recipient cells of various types and the extracellular matrix. For example, CODEX technology currently permits analysis of 40+ protein markers.

Sample collection and analysis

Sample collection can be envisioned at the time of chronic GVHD diagnosis, at a time that would be informative in assessing immediate drug effects, and then later to assess correlations with clinical response. Not all Tier 1 studies need be performed in the same institution or in real-time, but in the long run, such information could add substantially to our understanding of disease pathogenesis. Material storage in a biorepository facilitating a communal effort to perform individual assays in standardized fashion by experts in such testing would enable quality control assessment, avoid the need for real-time assays, reduce costs, minimize sample to sample variations if run in batches for a given patient, and ensure that the appropriate level of expertise is available for implementation and interpretation (see also below for sample collection and processing). A strong infrastructure at each site is needed with procedures for acquisition of cells in large enough numbers, plasma/serum and tissues, sample processing, storage, together with a governance process to distribute centrally stored samples and clinical data. As data are collected, some tests will prove more valuable than others, and the number of tests could be reduced,

retaining those that are most informative for understanding GVHD pathogenesis and the effects of specific interventions, in concert with the presence or absence of clinical response in each phenotype of the disease.

Call to action

This committee sends a call to action to require biological sampling in order to speed progress in developing new chronic GVHD therapies based on pathogenesis and biological and clinical responses. We call for the FDA/EMA, corporate partners, and patient advocates to require biological sampling and analysis. We call upon granting agencies to consider adequate financial support for such studies as fundamentally important to the field and to patient wellbeing.

Initial Therapy

The goals of treatment for chronic GVHD are numerous: to reduce symptom burden, control objective manifestations and prevent the progression of disease activity, preserve function by preventing irreversible damage and the resulting impairment and disability, and ideally to improve survival and allow or accelerate the development of operational tolerance that would allow withdrawal of all systemic treatment. In addition, these benefits must be sustained until systemic treatment is no longer needed, and the treatment must provide high therapeutic index in which benefits outweigh side effects.

Clinical studies should be designed to mirror clinical practice to the extent possible, to facilitate accrual of the more common clinical scenarios, and to test the biological and clinical effects of the study drug in the absence of confounding by other concurrent treatment. Up until now, nearly all

clinical trials of initial treatment for chronic GVHD have tested a study drug given in conjunction with corticosteroids. The biological effects of systemic corticosteroids in chronic GVHD are not well understood and their use blunts the ability to assess the true impact of the study intervention on pathophysiological signaling pathways and other potential mechanisms. Initial systemic monotherapy of chronic GVHD is an optimal setting to investigate the clinical and biological impact of an individual therapeutic agent.

A key question is whether new therapeutic agents may be tested for initial treatment of moderate or severe chronic GVHD in the absence of concurrent treatment with high-dose corticosteroids or whether trials for this indication should use add-on designs that include corticosteroids at prednisone-equivalent doses of 0.5 to 1.0 mg/kg/day. Studies testing new agents in the absence of concurrent corticosteroid treatment offer several advantages (**Figure 2**). In this "corticosteroid-free" approach, participants would initially receive systemic monotherapy with an investigational agent, and corticosteroid treatment would begin only if clinical manifestations of chronic GVHD worsen at any time or do not improve within an appropriate prespecified time after starting treatment. Multiple considerations support this novel approach. First, chronic GVHD has an insidious onset in most patients and does not typically require urgent intervention. Second, corticosteroid treatment causes considerable toxicity, and its omission or significant reduction could be beneficial even if the onset of improvement is delayed. Third, concurrent corticosteroid treatment masks the treatment effect of the interventional agent, since clinical manifestations improve initially in most patients after starting corticosteroid treatment. Finally, corticosteroid treatment can alter biomarkers used to measure biological effects of the study intervention. Other approaches to concurrent corticosteroid therapy could lower starting doses (i.e., 0.25 mg/kg daily,

rather than 0.5-1 mg/kg daily) and shorter duration of exposure (i.e., taper off within 1-3 months after starting treatment).

Eligibility Criteria

Inclusion criteria should articulate the characteristics of patients deemed to require systemic treatment. Moderate or severe chronic GVHD by NIH criteria is a clear indication for systemic treatment. In addition, mild chronic GVHD by NIH criteria can be an indication for systemic treatment if prespecified high risk features such as progressive onset or low platelet count, among others, are present. Enrollment of patients with a broad spectrum of chronic GVHD manifestations will facilitate efforts to identify relationships between specific organ manifestations, clinical phenotypes and laboratory correlates. Administration of medications used to prevent or treat acute GVHD should be continued for monotherapy trials, but treatment regimens must have been stable for at least 2 weeks before beginning any study treatment. Finally, special considerations should be made for children, since the characteristics of immune reconstitution and chronic GVHD evolution in children differ from those in adults.[14]

Clinical Endpoints

The primary endpoints in phase I studies of "corticosteroid-free" initial monotherapy for chronic GVHD should be safety and feasibility. A clinical trial design that investigates multiple dose levels will help identify the ideal dose for subsequent trials, based on clinical effect, biological impact, and safety profile. Safety considerations include defined stopping rules for both futility and unacceptable toxicity. The primary feasibility endpoint would assess the proportion of patients who remain corticosteroid naïve at a specified time point, such as 4 to 8 weeks after initiation of

treatment. Participants must be given corticosteroids or other effective treatment at the onset of any definitive evidence of disease progression by any measure, even if the change does not increase overall disease severity by NIH grading criteria. Likewise, corticosteroids or other effective treatment must be started if no improvement is observed within a prespecified time after starting treatment with the investigational agent. The onset of corticosteroid treatment must be documented, including the clinical indications, dosing, and duration.

Primary efficacy endpoints in Phase IIa studies are acceptable, particularly for interventions with an established safety profile in patients with chronic GVHD. Efficacy endpoints that incorporate organ specific and overall NIH Consensus Response Criteria[15] at reasonable time points are preferred. Secondary clinical endpoints can incorporate details of response (such as time to response, durability, organ-specific response), patient-reported outcomes, and survival endpoints. Time to expected response, based on chronic GVHD manifestations, is important to consider. Cutaneous erythema, oral manifestations, transaminase elevation and diarrhea would be expected to improve within weeks after starting effective treatment, whereas measurable improvement of cutaneous sclerosis, fasciitis and joint disease would take much longer, and manifestations of damage such as oral and ocular sicca, bronchiolitis and vitiligo would not be expected to improve, although progression could be halted by effective treatment. Previous studies have shown that the median time interval from onset of systemic treatment to permanent withdrawal of all systemic treatment in patients with chronic GVHD exceeds 2 years. Initial responses measured by NIH criteria at 3 or 6 months after starting treatment offer little confidence that they will endure until treatment can be withdrawn. For this reason, phase II studies should follow participants for long-term outcomes (i.e., at least a year) to ensure that responses are truly durable.[16]

Safety Considerations

Initial therapy of chronic GVHD without concurrent corticosteroid treatment differs from the long-established standard of using corticosteroids for initial treatment in all patients. For this reason, providers may hesitate to offer monotherapy trials with investigational agents for certain patients, especially if clinical manifestations of chronic GVHD show rapid evolution from poorly controlled acute GVHD. Such patients may be enrolled in these trials, knowing that corticosteroid treatment can be started at any time at the discretion of the provider. A similar approach has already been used in testing sirolimus for treatment of acute GVHD,[17] and monotherapy trials testing itacitinib for acute GVHD (NCT03721965) and ibrutinib for chronic GVHD (NCT04294641) are in progress. For optimal protection of participant safety, protocols should require expedited reporting of corticosteroid treatment so that stopping rules can be implemented promptly according to prespecified criteria. A hypothetical single-arm phase II efficacy study design of 'corticosteroid-free' monotherapy with an efficacy measure that reflects benchmarked potential clinical benefit is included as a **Supplement**.

Beyond Initial Therapy

Even as corticosteroid-free trials for initial treatment advance, trials for steroid-refractory and treatment-refractory chronic GVHD will be needed. Opportunities to elucidate the biological implications of novel agents are more difficult in these settings because of heterogeneity in disease manifestations and treatment history. Nevertheless, well-designed studies could broaden the understanding of interactions between clinical disease, therapeutic response, and disease biology (**Figure 3**). The treatment-refractory setting is also an appropriate entry point to explore the safety

profile of agents which have not been previously tested in allogeneic hematopoietic cell transplantation or GVHD.[9] Based on lessons from past clinical trials, relevant knowledge gaps are addressed in the following categories: corticosteroid considerations, disease considerations, clinical endpoints and safety considerations.

Corticosteroid Considerations

Refining eligibility criteria in studies for steroid-refractory chronic GVHD can help identify more homogeneous populations in which the clinical and biological impact of therapy could be better evaluated and measured. First, a key factor to consider is the dose and duration of prior corticosteroid therapy, as it can be much more difficult to detect a clinical response in patients who have been heavily treated for longer periods of time. Definitions of steroid refractory, intolerant and dependent disease are currently used in studies that define the minimum criteria for steroid exposure and response, without providing an upper limit.[15, 18, 19] Studying patients early after the disease is determined to be steroid-refractory is recommended as a way of increasing the likelihood of observing a clinical response or detecting a biologic effect. The same issues apply to the use of other systemic agents, either sequentially or concurrently, in patients who have been treated with multiple lines of therapy. Second, the distinction between **steroid-refractory** and **treatment-refractory** should be clearly stated, as these two groups may be biologically and clinically different, leading to the possibility causing erroneous dismissal of an effective agent. The 2014 Clinical Trial Working Group Report suggested criteria for defining steroid-refractory chronic GVHD,[3] and recent efforts have been made to provide new definitions for refractory acute GVHD following the FDA approval of ruxolitinib.[20, 21]

Finally, how to best handle corticosteroid therapy at the time of enrollment into a clinical trial after steroid failure remains an open question. As new therapies with a higher potential for efficacy and a steroid-sparing effect enter the treatment landscape, it becomes increasingly important to address this issue in the study design. Should corticosteroids be tapered, discontinued, maintained or even increased at the beginning of therapy? Should short courses of pulsed corticosteroids be allowed? Could specific dosing regimens lead to a reduction in the cumulative dose administered over time? These key considerations should be included in clinical trial designs.

Disease Considerations

The manifestations of chronic GVHD change over time, usually becoming more difficult to reverse due to progressive organ involvement and the fibroproliferative nature of more advanced stages of the disease. Given the often-refractory nature of chronic GVHD beyond the front-line therapy, this setting opens opportunities to explore combinations of novel targeted agents with non-overlapping toxicities, to increase efficacy and safety. We recommend prioritizing clinical trials that enroll patients with earlier stages of the disease, in order to maximize the chance of observing responses to therapy and to facilitate a more accurate assessment of that response. Disease-related characteristics such as inflammatory or fibrotic manifestations are another important factor to consider in defining eligibility. These clinical factors could be used to create cohorts of subjects with a common clinical phenotype in which a specific biological signal of disease or treatment effect could be detected. While patients can often have multiple concurrent features of chronic GVHD, trials with cohorts of patients who share specific manifestations would potentially allow the treatment of steroid-refractory chronic GVHD to move towards a more personalized approach based on the clinical presentation and biological profile.

Clinical Endpoints

The definition of clinically meaningful endpoints in steroid-refractory or treatment-refractory chronic GVHD studies remains challenging and should incorporate goals of therapy in order to maximize benefit to participants. Other factors that determine clinical endpoints are the type of intervention, study phase, indication (i.e., steroid- or treatment-refractory) and safety considerations. The efficacy of any new chronic GVHD therapy beyond corticosteroids has 2 main components: response and steroid-sparing effect. Overall response rate (based on NIH Consensus Response Criteria) after 3 or 6 months of therapy has been used in many recent trials. Most responses in this setting are partial responses, and the clinical significance of a partial response can vary widely. As the partial response category is broad, additional measurements including patient-reported outcomes or biological correlates can be particularly useful. Time to response and durability of response beyond 6 months and their implications for additional therapies, steroid-sparing effect or survival outcomes are less clear, so these assessments should be incorporated whenever possible. The measurement of a steroid-sparing effect in advanced chronic GVHD is almost as relevant as response in defining efficacy, although steroid dose reduction per se does not qualify as evidence clinical benefit from a regulatory perspective. Although there are no standard metrics for the steroid-sparing effect, the magnitude, incidence, timing and durability of corticosteroid dose reduction should be documented. Whenever possible, the cumulative incidence of discontinuation of corticosteroids and durability should also be reported. As most of these studies are single-arm, the need for comparability with other clinical trials is highly relevant.

Safety Considerations

Evaluation of the safety and tolerability of novel agents in more heavily pretreated and fragile populations remains an important goal of future studies. As chronic GVHD is associated with high disease burden and morbidity over long periods of time, the potential benefit of therapy, including steroid-sparing potential, must outweigh the side effects and burden of treatment. In trials that predominantly enroll heavily pre-treated patients or those with more advanced chronic GVHD, safety stopping rules should be designed to incorporate relevant treatment-emergent adverse events that are not well tolerated.

Subsequent Clinical Trial Development

In the short term, therapeutic agents with clinical effectiveness in chronic GVHD will continue to emerge from ongoing and future clinical trials. Most of these studies will have been conducted as phase I or II trials, without a comparison arm. The committee recommends that these agents next be investigated in the context of randomized clinical trials. In upfront systemic treatment, two separate large phase III randomized clinical trials are investigating ibrutinib (NCT02959944) and itacitinib (NCT03584516) compared to placebo, in combination with corticosteroids for moderate or severe chronic GVHD. If a novel agent demonstrates safety and encouraging clinical efficacy in a "corticosteroid-free" study design described above, a follow up study could randomize subjects to treatment with the investigational agent versus standard of care corticosteroids. For patients with treatment-refractory chronic GVHD, a trial could randomize participants to treatment with the investigational agent versus best available therapy, as has been done in the investigation of ruxolitinib for the treatment of steroid-refractory acute GVHD (NCT02913261)[22] and chronic GVHD (NCT03112603). Ultimately, as more agents receive FDA approval for the treatment

refractory chronic GVHD, trials that randomize subjects between therapeutic agents will be needed. All randomized trials should continue to include correlative laboratory measures, to evaluate differences in baseline biomarker profiles and longitudinal changes after treatment. In the long term, multiple clinical trials with extensive correlative studies will begin to elucidate the relationships between biological and clinical responses.

As data aggregate in subsequent trials and correlation with clinical manifestations strengthen, master protocols may eventually emerge as a subsequent path for further investigation (**Figure 4**). Master protocols allow for multiple therapeutic agents or multiple disease phenotypes to be evaluated within the context of a single overall protocol, in attempt to advance the field quickly and efficiently.[23] In the context of chronic GVHD, a platform trial design could be considered, in which therapeutic agents with predictive probability of being more effective than standard therapy (i.e., corticosteroids for initial treatment) graduate from the trial with their corresponding biomarker signature, and agents are dropped if they show a low probability of improved efficacy. New agents enter the study to replace those that have completed the evaluation. The platform design also allows for standardization of study conduct, endpoints, and correlative study panels.[24] The most successful platform trial to date is the I-SPY 2 protocol, which investigates novel neoadjuvant therapies for early-stage breast cancer based on specific molecular subtypes of the disease.[25] A similar protocol could establish a collaborative space to accelerate the testing of novel agents in the treatment of chronic GVHD, but would require definition of specific clinical or molecular disease subtypes and multicenter institutional infrastructure and subsequent collaboration with individual companies or institutions to tests promising agents. While master protocols are not considered a short-term goal in chronic GVHD due to logistical considerations,

innovative approaches to clinical trial design could expedite testing of therapeutic agents for an important unmet clinical need.

Conclusions

Treatment options for patients with chronic GVHD have expanded in recent years, led by unprecedented large multi-center randomized trials through collaboration with industry and an FDA-approved treatment agent. Despite these advances, the understanding of interactions between clinical chronic GVHD variables and targeted therapies with the pathophysiologic pathways that have been identified in pre-clinical models remains very limited. Efforts should continue to prioritize the assessment of biological correlatives in the context of clinical trial designs that will maximize the ability to close these knowledge gaps and eventually redefine the paradigms of chronic GVHD into a more personalized approach to treatment. Within the next 3 years, clinical trials should challenge the current standard of care by investigating novel agents in a corticosteroid-free approach to initial systemic monotherapy. Novel agents demonstrating meaningful clinical response should then be tested along with other extensively evaluated therapeutic agents in the context of randomized clinical trials. As a less immediate goal, advanced collaboration between academic medical centers, medical societies, and industry could develop larger trials with innovative designs that enable rapid testing of novel agents with standardized clinical and biological evaluations.

ACKNOWLEDGMENTS

This project was supported by the National Institutes of Health (NIH): National Cancer Institute (NCI), Center for Cancer Research, Intramural Research Program and Division of Cancer Treatment and Diagnosis, Cancer Therapy Evaluation Program; Office of Rare Disease Research (ORD), National Center for Advancing Translational Sciences; Division of Allergy, Immunology and Transplantation, National Institute of Allergy and Infectious Diseases; National Heart Lung and Blood Institute, Division of Blood Diseases and Resources. The authors acknowledge the following 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 Deutsche José Carreras Leukämie-Stiftung. The organizers are indebted to patients and patient and research advocacy groups, who made this process much more meaningful by their engagement. 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. Project participants also recognize Dr. x, Dr. x, and Dr. x for their independent expert reviews and comments on documents prior to the November 2020 meeting.

Disclaimer: 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.

Conflict of interest statement:

REFERENCES

1. Socie G, Ritz J. Current issues in chronic graft-versus-host disease. *Blood* 2014; **124**(3)**:** 374-384.

2. 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**(12)**:** 2340-2346.

3. Martin PJ, Lee SJ, Przepiorka D, Horowitz MM, Koreth J, Vogelsang GB *et al.* National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: VI. The 2014 Clinical Trial Design Working Group Report. *Biol Blood Marrow Transplant* 2015; **21**(8)**:** 1343-1359.

4. Cutler CS, Koreth J, Ritz J. Mechanistic approaches for the prevention and treatment of chronic GVHD. *Blood* 2017; **129**(1)**:** 22-29.

5. Zeiser R, Blazar BR. Pathophysiology of Chronic Graft-versus-Host Disease and Therapeutic Targets. *The New England journal of medicine* 2017; **377**(26)**:** 2565-2579.

6. Miklos D, Cutler CS, Arora M, Waller EK, Jagasia M, Pusic I *et al.* Ibrutinib for chronic graft-versus-host disease after failure of prior therapy. *Blood* 2017; **130**(21)**:** 2243-2250.

7. Jaglowski SM, Blazar BR. How ibrutinib, a B-cell malignancy drug, became an FDA-approved second-line therapy for steroid-resistant chronic GVHD. *Blood Adv* 2018; **2**(15)**:** 2012-2019.

8. Dignan FL, Amrolia P, Clark A, Cornish J, Jackson G, Mahendra P *et al.* Diagnosis and management of chronic graft-versus-host disease. *British journal of haematology* 2012; **158**(1)**:** 46-61.

9. Im A, Hakim FT, Pavletic SZ. Novel targets in the treatment of chronic graft-versus-host disease. *Leukemia* 2017; **31**(3)**:** 543-554.

10. 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**(2)**:** 126-137.

11. Gandelman JS, Byrne MT, Mistry AM, Polikowsky HG, Diggins KE, Chen H *et al.* Machine learning reveals chronic graft-versus-host disease phenotypes and stratifies survival after stem cell transplant for hematologic malignancies. *Haematologica* 2019; **104**(1)**:** 189-196.

12. Franks JM, Martyanov V, Cai G, Wang Y, Li Z, Wood TA *et al.* A Machine Learning Classifier for Assigning Individual Patients With Systemic Sclerosis to Intrinsic Molecular Subsets. *Arthritis Rheumatol* 2019; **71**(10)**:** 1701-1710.

13. Paczesny S, Hakim FT, Pidala J, Cooke KR, Lathrop J, Griffith LM *et al.* National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host

Disease: III. The 2014 Biomarker Working Group Report. *Biol Blood Marrow Transplant* 2015; **21**(5)**:** 780-792.

14. Cuvelier GDE, Nemecek ER, Wahlstrom JT, Kitko CL, Lewis VA, Schechter T *et al.* Benefits and challenges with diagnosing chronic and late acute GVHD in children using the NIH consensus criteria. *Blood* 2019; **134**(3)**:** 304-316.

15. Lee SJ, Wolff D, Kitko C, Koreth J, Inamoto Y, Jagasia M *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. The 2014 Response Criteria Working Group report. *Biol Blood Marrow Transplant* 2015; **21**(6)**:** 984-999.

16. Martin PJ, Storer BE, Inamoto Y, Flowers MED, Carpenter PA, Pidala J *et al.* An endpoint associated with clinical benefit after initial treatment of chronic graft-versus-host disease. *Blood* 2017; **130**(3)**:** 360-367.

17. Pidala J, Hamadani M, Dawson P, Martens M, Alousi AM, Jagasia M *et al.* Randomized multicenter trial of sirolimus vs prednisone as initial therapy for standard-risk acute GVHD: the BMT CTN 1501 trial. *Blood* 2020; **135**(2)**:** 97-107.

18. Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW *et al.* 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. *Biol Blood Marrow Transplant* 2015; **21**(3)**:** 389-401 e381.

19. Schoemans HM, Lee SJ, Ferrara JL, Wolff D, Levine JE, Schultz KR *et al.* EBMT-NIH-CIBMTR Task Force position statement on standardized terminology & guidance for graft-versus-host disease assessment. *Bone marrow transplantation* 2018; **53**(11)**:** 1401-1415.

20. Zeiser R, Socie G. The development of ruxolitinib for glucocorticoid-refractory acute graft-versus-host disease. *Blood Adv* 2020; **4**(15)**:** 3789-3794.

21. Mohty M, Holler E, Jagasia M, Jenq R, Malard F, Martin P *et al.* Refractory acute graft-versus-host disease: a new working definition beyond corticosteroid refractoriness. *Blood* 2020; **136**(17)**:** 1903-1906.

22. Zeiser R, von Bubnoff N, Butler J, Mohty M, Niederwieser D, Or R *et al.* Ruxolitinib for Glucocorticoid-Refractory Acute Graft-versus-Host Disease. *The New England journal of medicine* 2020; **382**(19)**:** 1800-1810.

23. Woodcock J, LaVange LM. Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both. *The New England journal of medicine* 2017; **377**(1)**:** 62-70.

24. Berry SM, Connor JT, Lewis RJ. The platform trial: an efficient strategy for evaluating multiple treatments. *Jama* 2015; **313**(16)**:** 1619-1620.

25. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clinical pharmacology and therapeutics* 2009; **86**(1)**:** 97-100.

FIGURES AND LEGENDS

Figure 1. Exploring the associations between clinical disease manifestation, biological profiles, and treatment effect in chronic GVHD.

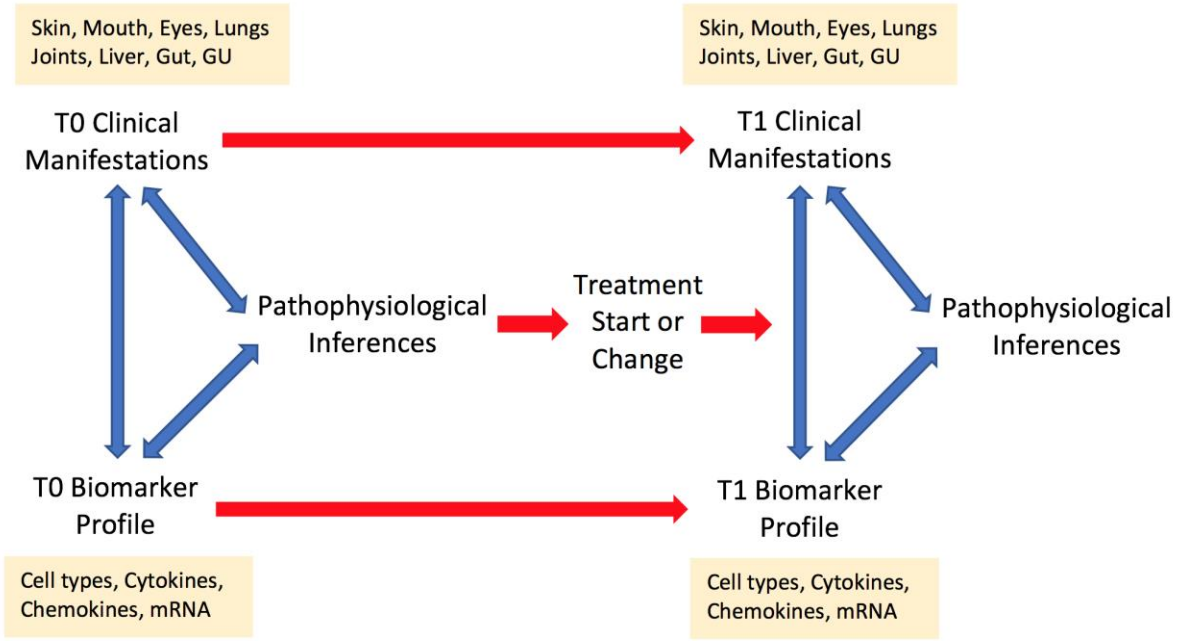


Figure 2. Proposed framework and considerations of front-line treatment clinical trials in chronic GVHD.

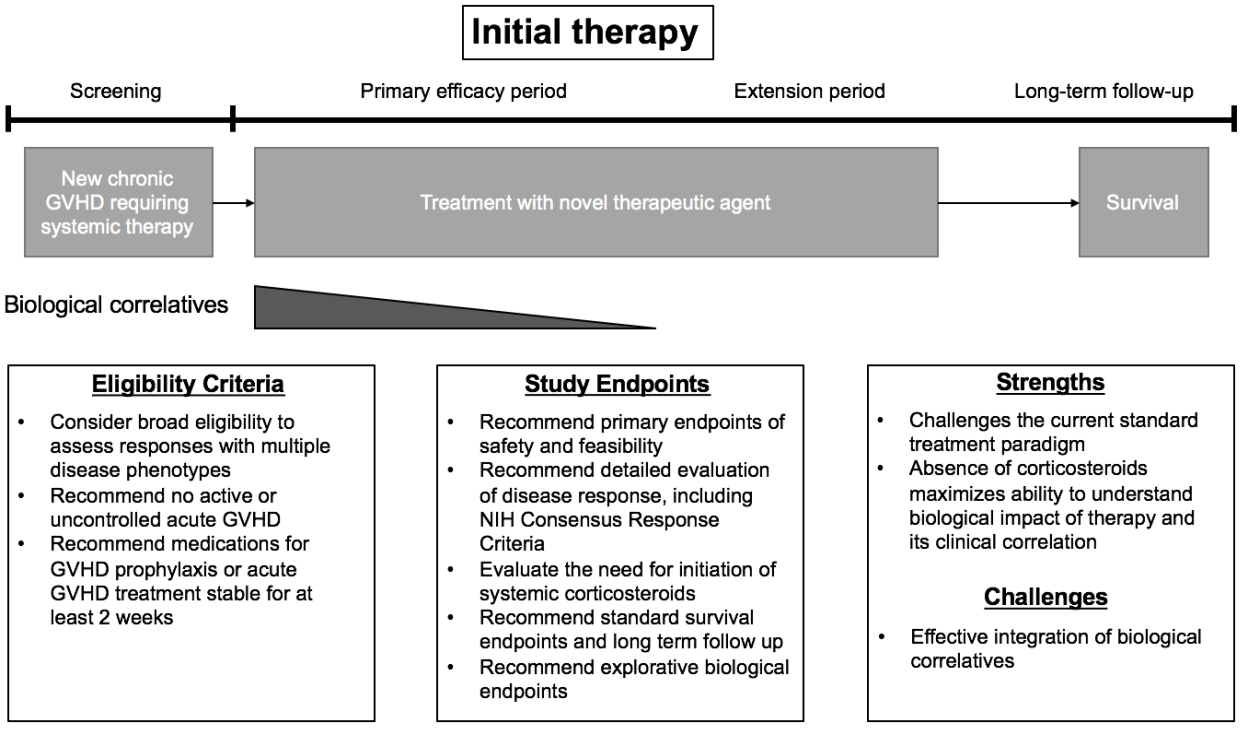
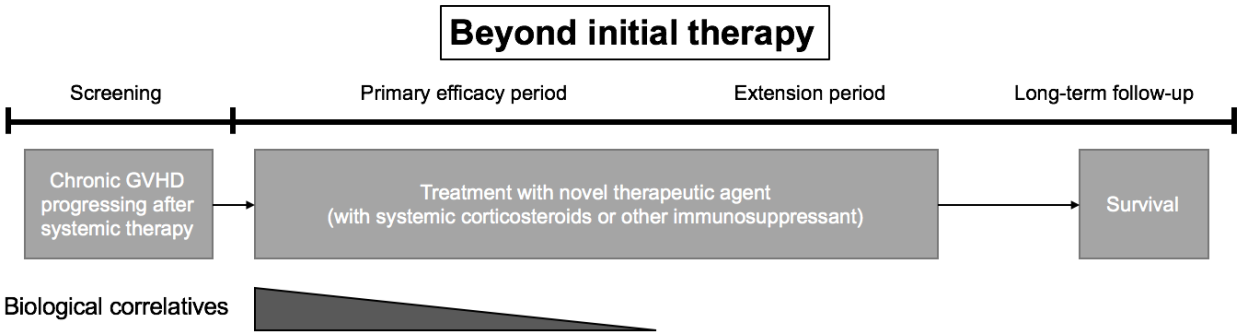


Figure 3. Proposed framework and considerations of second-line treatment clinical trials in chronic GVHD.



Eligibility Criteria

- Consider subjects with shorter duration of disease
- Consider enrollment of subjects with common disease manifestations or phenotypes
- Consider limitations on eligibility based on corticosteroid exposure (duration, dose, etc.)

Study Endpoints

- Recommend detailed evaluation of disease response, including NIH Consensus Response Criteria
- Emphasize importance of evaluating disability symptoms and quality of life
- Recommend explorative biological endpoints

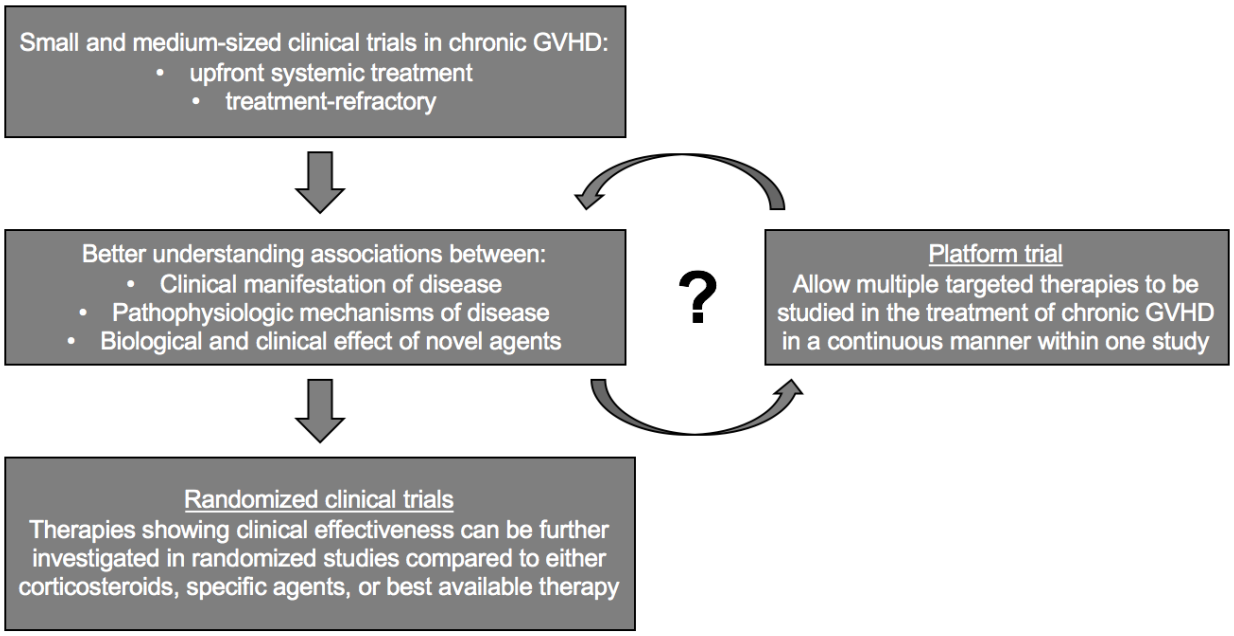
Strengths

- Evaluates the biological and clinical impact of novel agents in pretreated population

Challenges

- Creating clinically meaningful eligibility criteria in a heterogeneous population
- Effective integration of biological correlatives

628 Figure 4. Future directions of chronic GVHD clinical trials.



629

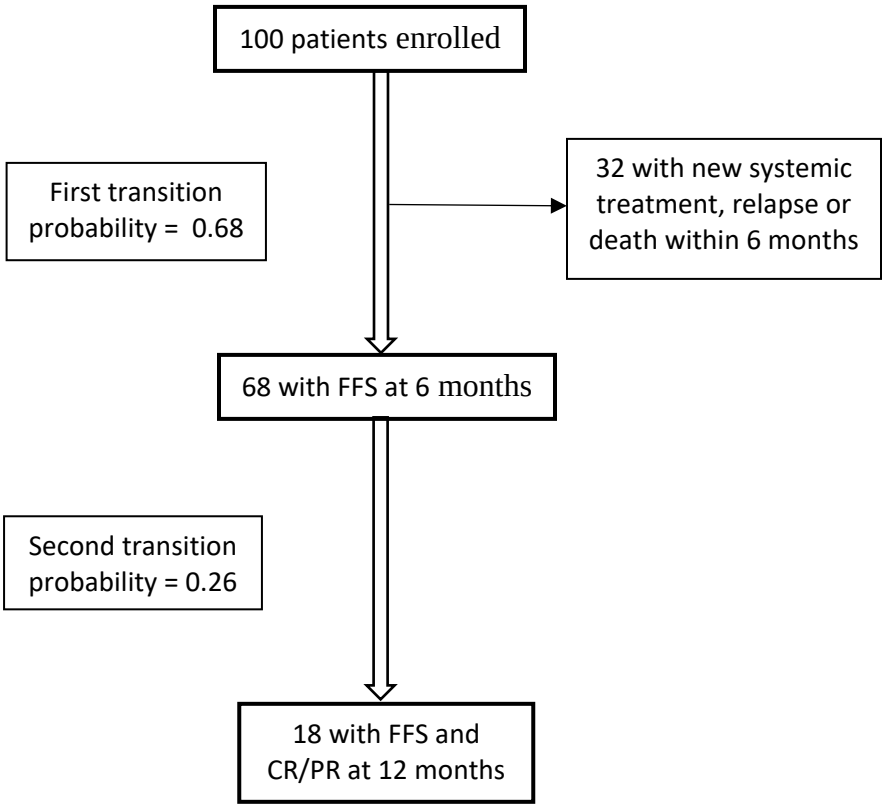
Supplement

Material in this supplement shows an example of how a single-arm phase II efficacy study of steroid-free initial treatment for chronic GVHD could be designed to incorporate stopping rules for safety, stopping rules for futility and a pre-specified efficacy measure that reflects benchmarked potential clinical benefit with a reasonable number of patients.

Demonstration of success in testing an investigational product for treatment of chronic GVHD requires participant survival without recurrent malignancy or new systemic treatment before the endpoint assessment (i.e., failure-free survival, FFS) and a complete or partial response (CR/PR) at the time of the endpoint assessment. Results of a prospective observational study showed that the probability of FFS at 6 months after initial treatment was 68%.[1] Only 26% of the patients with FFS at 6 months had FFS with a CR/PR at 12 months. Therefore, 18% of the patients in the starting population had FFS with a CR/PR at 12 months (Figure), an endpoint associated with improved survival beyond 12 months. At 6 months, the proportion of patients with FFS and a CR/PR was 31%. The lower proportion of patients with FFS and a CR/PR at 12 months indicates that many of the 6-month outcomes were not durable.

The probability of FFS with CR/PR at 12 months is unlikely to improve unless an investigational product improves both the 68% transitional probability of FFS at 6 months and the 26% transitional probability of FFS with CR/PR among these patients at 12 months. For example, improvement in the 6-months FFS probability from 68% to 82% and improvement in the second transitional probability from 26% to 40% would improve the final success rate from 18% to 33%. Importantly, a trial could be stopped if the interim results showed that the goal for the first transitional probability could not be met. Under a null hypothesis of 0.68 and an alternative of 0.82 for the first transitional probability, a sample size of 44 patients would be needed for a trial with 80% power and a 1-sided type-1 error of 0.10. Success for this interim endpoint would require FFS at 6 months in at least 35 of the 44 patients. Failure to meet this endpoint in any 10 patients would indicate that treatment with the investigational product is not likely to improve the final success rate and would warrant termination of the trial. Under a null hypothesis of 0.18 and an alternative of 0.33 for the final success rate, a trial enrolling 44 patients would have 86% power with a 1-sided type-1 error of 0.1. The high one-sided type-1 error would be acceptable, because the results are intended only to indicate the merits of conducting a pivotal trial and to inform its design.

1. Martin PJ, Storer BE, Inamoto Y, Flowers MED, Carpenter PA, Pidala J *et al.* An endpoint associated with clinical benefit after initial treatment of chronic graft-versus-host disease. *Blood* 2017; **130**(3)**:** 360-367.



Flow of outcomes in a multi-center prospective observational study of initial treatment for chronic GVHD.[1] The study included 328 patients, but the figure is adapted to reflect outcomes as estimated proportions of the total. Results of this study have not been tested for replication in other cohorts. FFS indicates failure-free survival and CR/PR indicate complete or partial response.