Current Hematologic Malignancy Reports (2019) 14:22–30
https://doi.org/10.1007/s11899-019-0494-x

STEM CELL TRANSPLANTATION (R MAZIARZ, SECTION EDITOR)

Graft Versus Host Disease Clinical Trials: Is it Time for Patients Centered Outcomes to Be the Primary Objective?

Bronwen E. Shaw[1]

Published online: 14 January 2019

Abstract

Purpose of Review Graft versus host disease (GVHD) is a common complication following hematopoietic cell transplant and is associated with a high symptom burden, reduced functional status, and impaired quality of life (QOL). QOL is best assessed by patient-reported outcomes (PRO). Numerous clinical trials for the prevention and treatment of GVHD are available. This review aims to understand the landscape of PRO inclusion in clinical trials for GVHD over the last decade.

Recent Findings Consensus bodies, including experts in GVHD, PRO, and clinical trials have made recommendations for a standardized approach for the inclusion of PRO in clinical trials including as primary outcomes, however, these have yet to be implemented in a consistent manner in practice.

Summary Consistently applying consensus recommendation in chronic GVHD will ensure that PROs are appropriately included in clinical trials. Development of validated measures in acute GVHD and composite outcomes for all GVHD trials are required.

Keywords Hematopoietic cell transplantation · Graft versus host disease · Patient-reported outcomes · Clinical trials

Introduction

The field of hematopoietic cell transplantation (HCT) is increasingly embracing a patient-centered approach to care including the collection and assessment of patient-reported outcomes (PRO), subjective reports directly from the patient, describing their symptom burden, functional status, and overall quality of life (QOL). PROs are beginning to be incorporated into both clinical care and research, but their use is currently not standardized, whether considering the measures to use, the time points for collection, the mode of collection, and the interpretation and action which should result from the PRO scores. Likewise, their use in clinical trials frequently remains unstandardized and often PRO endpoints are not treated with the same rigor as other endpoints in the protocol [1].

Graft versus host disease (GVHD) is a very common complication following HCT, with numerous clinical trials addressing the use of novel agents or strategies for both prevention and treatment. Acute GVHD (aGVHD) predominantly affects the skin, liver, and GI tract and may be associated with very troublesome symptoms, physician report of some of which form the basis for grading and staging. Chronic GVHD (cGVHD) is a multisystem disease associated with an increased symptom burden, reduced functional status, and impaired QOL with increased mortality. Given this profile, GVHD is a disease where the patient's perspective is critical.

The aim of this review is to summarize the landscape of PRO inclusion in clinical trials in GVHD over the last decade, addressing (1) clinical trials that are currently ongoing or recently completed, (2) consensus statements by expert groups in GVHD, and (3) advances in PROs outside of GVHD but relevant to clinical trials, HCT, and GVHD.

PROs in GVHD Clinical Trials: Position in Late 1990s/Early 2000s

In 2009, Lee et al. published a consensus document reporting on discussion by several interested groups (including the Food and Drug Administration (FDA), National Institutes of Health

This article is part of the Topical Collection on *Stem Cell Transplantation*

✉ Bronwen E. Shaw
beshaw@mcw.edu

[1] CIBMTR (Center for International Blood and Marrow Transplant Research)/Froedtert, Department of Medicine, Medical College of Wisconsin, 9200 W. Wisconsin Ave., Suite C5500, Milwaukee, WI 53226, USA

(NIH), Center for International Blood and Marrow Transplantation Research (CIBMTR), and American Society for Blood and Marrow Transplantation (ASBMT)) on the role of PROs in trial design and interpretation for aGVHD prevention and treatment trials [2]. They concluded that GVHD prophylaxis studies are unlikely to benefit from PRO collection as the current agents to prevent GVHD are relatively ineffective, resulting in small changes in GVHD rates and that these agents are given at the time in early post-HCT when patients have many symptoms related to multiple causes [2]. For treatment trials in aGVHD, designing a measure and collecting PROs are challenging due to several factors including the need for frequent measures at a time when patients may be very unwell, certain characteristics of GVHD (e.g., the liver involvement) not resulting in symptoms, and many other post-HCT factors contributing to symptoms at this time [2]. Nevertheless, the authors conclude that there is an urgent need for a validated instrument that meets the requirement for a FDA claim approval for use in aGVHD treatment trials. In cGVHD, the 2006 NIH Design of Clinical Trials Working Group report recommended the inclusion of PROs in cGVHD clinical trials as secondary outcomes. Although the value of PROs is noted, they state that PROs are subject to problems of bias and missing data, may lack specificity and that few measures show evidence of sensitivity to change. No specific PRO measures were proposed. They call for development of pediatric measures [3].

FDA guidance on the requirements for a PRO instrument to be considered adequate as an endpoint, at least in clinical trials for medicinal product labeling, was published in 2009 [4]. This includes a requirement for careful consideration of the role of the endpoint in each trial, the need for the PRO instrument to be valid and reliable, including in the specific target patient population, and that the PRO outcome be supported by an objective clinical measure of drug activity [4, 5].

Current or Recent GVHD Clinical Trials

A search of clinicaltrials.gov using the term "Graft Versus Host Disease" reveals 457 studies (October 2018). Incorporating either "patient-reported outcomes" or "quality of life" identifies eight and 37 trials respectively. After accounting for duplicates ($n = 2$), withdrawn/unknown status ($n = 3$), not clinical trial ($n = 5$), outcomes not GVHD-specific ($n = 4$), and QOL/PRO not outcome ($n = 2$), a total of 29 studies remain. These can be divided broadly into the category of prophylaxis/prevention studies ($n = 4$, Table 1), GVHD treatments studies in the recruitment phase ($n = 9$, Table 2) and GVHD treatments studies which have completed ($n = 16$, Table 3). PRO measures, available on clinicaltrials.gov, are included in the tables. Studies include between one and five measures. There are seven aGVHD treatment trials, using the Functional Assessment of Cancer Therapy-Bone Marrow Transplant (FACT-BMT) [6] ($n = 2$), MD Anderson Symptom Inventory (MDASI) [7] ($n = 2$), short-form-36 (SF36) [8] ($n = 1$), and the EQ-5D [9] ($n = 1$). In the 18 cGVHD treatment trials, the Lee cGVHD Symptom Scale (Lee SS) [10] is most commonly used ($n = 8$), followed by the FACT-BMT ($n = 6$), SF36 ($n = 5$), and Human Activities Profile (HAP) [11] ($n = 4$). A variety of other measures are used less frequently.

The PRO/QOL assessment is listed as a secondary objective in all but two of the 29 trials. The first is a phase II GVHD prevention study in which the agent, Vorinostat is added to the standard GVHD prophylaxis of tacrolimus and methotrexate (NCT02409134) [12]. As Vorinostat, an HDAC inhibitor, has been shown to have neuroprotective and neurorestorative effects in preclinical models, the authors investigated the neurocognitive function and QOL in study patients using a variety of measures in a longitudinal manner. Control subjects received either an allogeneic (without Vorinostat) or autologous transplant. The results showed that patients receiving Vorinostat had a total neurocognitive performance comparable with autologous controls, who performed better than allogeneic controls. In this study, the incidence of GVHD was not reported between cases and controls. The second is the POSTAGE study, prospectively evaluating the outcomes of second-line therapies in aGVHD, which is in the recruitment phase (NCT02151539). QOL, as assessed by the FACT-BMT measure, is one of the four primary objectives.

Of the 16 completed studies (five were terminated due to slow accrual), eight (50%) have a publication of their results available, which in six (75%) cases included the PRO/QOL data (Table 3). Three unpublished trials were completed within the last 5 months (a publication may be pending). Of the six papers which included PRO outcomes, one was for initial therapy of acute GVHD, one for initial therapy of chronic GVHD, and four for steroid-refractory GVHD.

The Blood and Marrow Transplant Clinical Trials Network (BMT CTN) 0802 (NCT01002742) is a phase III, randomized double-blind, placebo-controlled trial evaluating the addition of mycophenolate mofetil (MMF) vs. placebo to systemic corticosteroids as initial therapy for aGVHD [13]. A pre-planned futility rule was met after 235 patients were enrolled, as MMF did not improve GVHD-free survival compared to steroids alone. The MDASI measure was completed for 103 MMF patients and 108 placebo patients, approximately 90% of patients. The total MDASI symptom severity or interference score was no different by arm.

BMT CTN 0801 (NCT01106833) is a phase II/III randomized, multicenter trial of prednisone/sirolimus vs prednisone/sirolimus/calcineurin inhibitor for the treatment of cGVHD at first presentation. The study was launched with three arms, but the prednisone/sirolimus/photophoresis-arm closed prematurely due to slow accrual and the remaining two-drug vs.

Table 1 GVHD prophylaxis studies listed on clinicaltrials.gov, which include quality of life (QOL) or patient-reported outcomes (PRO) endpoints

Study	GVHD type and endpoint	Study agent/ intervention	PRO primary/ secondary	Measure(s)	Study status	Publication	PRO reported in publication?
NCT02409134	GVHD prophylaxis	Vorinostat	Primary (feasibility)	Cogstate PHQ-9 GAD-7 FACT-G	Completed	Yes	Yes
NCT00391170	Chronic GVHD prophylaxis	Topical dexamethasone oral solution	secondary	OHIP-14	Recruiting		
NCT03339297	Acute GVHD prophylaxis	Defibrotide	Secondary	FACT-BMT-TOI EQ-5D	Recruiting		
NCT02345850	Acute and chronic GVHD prophylaxis	Calcineurin inhibitor-free interventions	Secondary	QOL survey	Recruiting		

Cogstate, Computerized Cognitive Test; *PHQ-9*, Patient Health Questionnaire; *GAD-7*, General Anxiety Disorder; *FACT-G*, Functional Assessment of Cancer Therapy-general; *OHIP-14*, 14-item Oral Health Impact Profile; *FACT-BMT*, Functional Assessment of Cancer Therapy-Bone Marrow Transplant; *TOI*, trial outcome index; QOL survey = measure not stated

three-drug study ended in phase II for statistical futility with 138 evaluable subjects [14]. QOL was measured by the FACT-BMT, SF-36, and Lee SS. Patients on the two-drug arm had significantly better scores at 2 and 6 months as measured by the SF36 physical component summary (adjusted by baseline score) than the three-drug arm. The FACT-BMT physical well-being scores were also higher at 2 months on the two-dug arm, but only in the non-adjusted analysis.

The final four studies investigated novel agents in steroid-refractory GVHD. Baird et al. (NCT00702689) published a pilot study investigating the impact of imatinib in patients with sclerotic skin changes. Twenty patients were enrolled, 14 of whom were evaluable for the primary endpoint of response at 6 months, defined by changes in their range of motion (ROM). Five had a PR and seven had stable disease. ROM was not significantly associated with any functional or PRO, including grip strength, walk times, the HAP, Lee SS, SF36, and cGVHD Activity Assessment-Patient Self Report [15]. Twenty-one patients were enrolled in phase I/II study investigating response to rituximab (NCT00136396) [16]. Seventy percent of available patients ($n = 20$) obtained clinical responses. Nine evaluable patients had an improvement in their Lee SS score, seven of whom had also been assessed as having a clinical response. Thalidomide (NCT00075023) was tested in phase II, randomized, placebo-controlled, double-blind trial evaluating its impact on patients with treatment-refractory oral cGVHD [17]. The study terminated early, with only ten subjects accrued. While several interesting observations were made (including some response to placebo), the small numbers precluded any firm statistical conclusions. Finally, sonidegib, a selective antagonist of the hedgehog coreceptor Smoothened, was tested in the phase I trial (NCT02086513). Accrual was terminated early due to toxicity burden (not attributed to sonidegib) and patient decisions to stop taking sonidegib. Approximately, half of the patients had a clinical response (PR) and there was a significant change in the Lee SS scores between responders and non-responders—although the mean change in score within each group was not significant. The FACT-G showed no change between responders and non-responders [18]. PRO completion rates are not uniformly available.

PROs in Acute GVHD: Progress Since 2010

It is not clear that significant progress has been made in the development of a GVHD-specific PRO measure and incorporating PRO in aGVHD clinical trials, although investigators have called for the development of novel endpoints such as PROs in studies [19]. Recent international consensus publications by Harris et al. [20] and Schoemans et al. [21] make significant strides in the standardization of terminology, diagnosis, and disease assessment in aGVHD, and while the goal was not specifically to address clinical trials or PROs, a firm foundation is laid to do so.

PROs in Chronic GVHD: Progress Since 2010

Significant advances have, however, been made in the understanding of how to incorporate PROs into cGVHD clinical trials. Although not a clinical trial, the chronic GVHD consortium performed a prospective multicenter observational study in a cohort of HCT recipients requiring immunosuppression for cGVHD. Patients were followed every 6 months collecting both physician- and patient-reported standardized information on organ involvement and symptoms. NIH consensus scoring was reported by clinicians, and PRO were collected using the SF36, FACT-BMT, Lee SS, Human activity profile, activity scale for kids, and the ocular surface disease index [22]. An aim of the PRO data collection was to evaluate the measures to

Table 2 GVHD treatment studies listed on clinicaltrials.gov, which include quality of life (QOL) or patient-reported outcomes (PRO) in recruitment

Study	Acute/chronic GVHD	Study intervention	PRO primary/ secondary	Measure(s)	Study status
NCT02151539	Acute, second-line therapy	Therapies including extracorporeal photophoresis	Primary (one of)	FACT-BMT	Recruiting
NCT02913261	Acute steroid refractory	Ruxolitinib	Secondary	FACT-BMT EuroQol-5D-5 L	Recruiting
NCT02652130	Acute steroid refractory	Remestemcel-L (mesenchymal stromal cell)	Secondary	QOL survey	Recruiting
NCT03640481	Chronic	Efficacy and safety of KD025	Secondary	Lee SS	Not yet recruiting
NCT02669251	Chronic (bronchiolitis obliterans syndrome)	AZD9668, an oral neutrophil elastase inhibitor	Secondary	Lee SS HAP FACT-BMT 6 min walk test	Recruiting
NCT01273207	Chronic (bronchiolitis obliterans syndrome)	Cyclosporine inhalation solution (CIS)	Secondary	QOL survey	Recruiting
NCT03007238	Chronic, steroid refractory	Extracorporeal photophoresis and low-dose aldesleukin	Secondary	Chronic GVHD Symptom Scale	Recruiting
NCT03422627	Chronic, steroid refractory	AMG 592 (IL-2 mutein)	Secondary	SF36 Lee SS	Recruiting
NCT03112603	Chronic, steroid refractory	Ruxolitinib	Secondary	Lee SS FACT-BMT EQ-5D	Recruiting

FACT-BMT, Functional Assessment of Cancer Therapy-Bone Marrow Transplant; *Lee SS*, Lee cGVHD Symptom Scale (Lee SS); *HAP*, Human Activities Profile; *SF36*, Short-form-36; QOL survey = measure not stated

determine which would be most useful in evaluating GVHD symptom burden in clinical trials. Analysis of PRO at baseline (a single time point) showed significant differences between the PRO scores in patients with different NIH disease severity (mild, moderate, severe) and concluded that there is a significant association between PROs and NIH criteria even when adjusting for other important transplant characteristics. The SF36 and FACT-BMT performed equally well in this cross-sectional setting [23••]. A separate analysis addressed changes in GVHD activity over time. They found that changes in patient-reported cGVHD were significantly associated with changes in PROs, that clinician-assessed cGVHD changes were associated with changes in PROs, but that changes in the NIH-assessed global cGVHD were not associated with changes in PROs. They also reported that a smaller set of questions from the FACT-BMT, the FACT-G, was as good as the full measure or the SF36 and could be used alone to reduce respondent burden; neither, however, correlated well with the NIH global severity scores [24].

Parallel efforts have been made to develop a PRO measure, specifically validated for use in cGVHD, which is able to assess change in the severity of cGVHD. The Lee Chronic GVHD Symptom Scale [10], developed with patient input in patients with active cGVHD, is a reliable and valid measure which been shown to be sensitive to change. The LSS is a 30-item 7-domain symptom scale, including the skin, eyes, and mouth; breathing, eating, and digestion; muscles and joints; energy and mental and emotional, which takes a median of 2 min to complete. Studies have shown that changes in the NIH skin, eye, mouth, GI, lung and summary scales correlate with patient-reported changes in GVHD [25–29]. Recently, further work has been done to re-examine the validity of the LSS in a contemporary patient cohort, towards FDA qualification of this PRO measure as a drug development tool. The results show very positive responses from participants, who reported that the scale is clear and captures almost all of their symptoms [30]. Although future studies may be needed to expand the populations studied, this study strongly supports the FDA qualification of this PRO measure.

In 2015, the NIH Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: VI. Design of Clinical Trials Working Group report [31] [Martin, 2015] was published. This report focuses on considerations for the design of cGVHD treatment (not prophylaxis) clinical trials and expands and improves upon the 2006 WG recommendations [3] by incorporating important advances in diagnosis and staging, biomarkers, histopathology, supportive care, and PRO studies (such as those mentioned above) that occurred over the intervening decade. The 2015 report is focused primarily on the development, characterization, validation, and selection of primary and secondary endpoints for studies of GVHD treatment. Although the considerations for selecting endpoints described in the 2015 report clearly address patient-centered outcomes "…..endpoints should be selected for their ability to demonstrate clinical benefit, which can be a prolongation of survival or an

Table 3 GVHD treatment studies listed on clinicaltrials.gov, which include quality of life (QOL) or patient-reported outcomes (PRO): completed

Study	Acute/ chronic	Study intervention	PRO primary/ secondary	Measure(s)	Study status	Publication	PRO reported in publication
NCT00929695	Acute, initial treatment	Low-dose glucocorticoids	Secondary	MDASI	Completed (December 2015)	No	
NCT02411084	Acute, steroid refractory	BEGEDINA®	Secondary	SF36	Terminated (insufficient rate of accrual)	No	
NCT01530256	Acute, steroid refractory	ALD518 (anti-interleukin 6 monoclonal antibody)	Secondary	FACT-BMT	Terminated (March 2013)	No	
NCT01002742	Acute	Steroids/mycophenolate mofetil vs steroids/placebo	Secondary	MDASI	completed	Yes	Yes
NCT02491359	Chronic	Carfilzomib	Secondary	Lee SS, HAP, SF36 FACT-BMT 2-min walk test	Completed (September 2018)	No	
NCT00702689	Chronic skin	Imatinib mesylate	Secondary	HAP SF36, Lee SS	Completed	Yes	Yes
NCT00136396	Chronic, steroid refractory	Rituximab	Secondary	QOL survey	Completed	Yes	Yes
NCT02123966	Chronic, steroid refractory	Topical sirolimus	Secondary	OHIP-14	Terminated (slow accrual)	No	
NCT00075023	Chronic	Topical thalidomide	Secondary	QOL survey	Terminated (unable to enroll)	Yes	Yes
NCT01287078	Chronic (Bronchiolitis obliterans syndrome)	Cyclosporine inhalation solution (CIS)	Secondary	QOL survey	Completed (August 2018)	No	
NCT02086513	Chronic, steroid refractory	LDE225 (Hedgehog signaling pathway inhibitor)	Secondary	Lee SS FACT-G	Terminated	Yes	Yes
NCT00031824	Chronic	Hydroxychloroquine	Secondary	QOL survey	Completed	Yes	No
NCT00144430	Chronic	Pentostatin	Secondary	QOL survey	Completed	Yes	No
NCT01380535	Chronic	Extracorporeal photophoresis (ECP) Therapy	Secondary	SF36 FACT-BMT	Completed (March 2017)	No	
NCT02513498	Chronic	Ixazomib citrate	Secondary	Lee SS HAP SF36 FACT-BMT	Completed (June 2018)	No	
NCT01106833	Chronic	Sirolimus plus prednisone and sirolimus/calcineurin inhibitor plus prednisone	Secondary	SF36 FACT-BMT	Active, not recruiting	Yes	Yes

MDASI, MD Anderson Symptom Inventory; *SF36*, Short-form-36; *FACT-BMT*, Functional Assessment of Cancer Therapy-Bone Marrow Transplant; *Lee SS*, Lee cGVHD Symptom Scale; *HAP*, Human Activities Profile; QOL survey = measure not stated; *OHIP-14*, 14-item Oral Health Impact Profile; *FACT-G*, Functional Assessment of Cancer Therapy-general; QOL survey = measure not stated

improvement in the way a patient feels or functions," there remain barriers to selecting PRO as primary endpoints, including the stringent FDA requirements mentioned above [4]. Thus, while the Consensus Committee makes a stronger recommendation about the inclusion of PROs than was made in 2006, it may still be considered soft, "patient-reported measures should be incorporated whenever feasible."

Conversely, the measuring therapeutic response in chronic graft-versus-host disease: NIH Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: IV. Response Criteria Working Group report [32••], makes a strong recommendation for the inclusion of a PRO, the LSS, as a "chronic GVHD-specific core measure," alongside clinician-assessed and patient-reported

signs and symptoms and clinician-assessed or patient-reported global rating scales. Non-GVHD-specific PRO measures such as the SF36, FACT-BMT, or HAP are strongly encouraged, but considered optional and should not be used as primary endpoints due to the lack of evidence of sensitivity to change for this indication [32••].

An important consideration for all data in clinical trials, including PROs, is completeness and missing data. A rigorous plan, outlined in the protocol and logistically in place from study onset, to ensure timely completion of PRO with regular reminders to participants and oversight by the study team, is required. The prospectively collected PRO data in the Chronic GVHD Consortium mentioned above had 87% completeness at baseline [23••], and completeness at follow up time points ranging from 75 to 85% [33, 34]. Statistical approaches to handle missing data should be included in the protocol.

Progress in the Use of PROs in Clinical Trials

Guidelines for the Inclusion and Reporting of PROs in Clinical Trials

Even for clinical trials that do include a PRO endpoint, it is not uncommon to find inadequate detail in the protocol, for example on background, methodology, hypothesis, and power [35] or sub-optimal reporting of the PRO outcomes in the study publication [36]. A recent review [1] of almost 800 clinical trials abstracted data related to eight outcomes from recommended quality standard for reporting PRO. They found that the proportion of clinical trials who met the eight indicators ranged between 15 and 81%. This inadequate design undermines not only the ability to interpret the PRO in the context of the study but also the confidence of the community in the value of the endpoint. The Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement was published in 2013 to standardize the quality of clinical trial protocols [37, 38]. An extension, providing specific recommendations for items to be addressed when PROs are included as either primary or secondary endpoints, was recently published [39]. This provides a number of extensions and elaborations on the checklist provided in the initial publication. The Consolidated Standards of Reporting Trials (CONSORT) statement provides evidence-based recommendations to improve the completeness of reporting of randomized controlled trials (RCTs). The most recent CONSORT Statement, published in 2010 [40], did not specifically address PRO, however, based on findings such as those mentioned above, a CONSORT-PRO extension was published in 2013 [41, 42], which describes the methodology by which the extension was developed and urges that all aspect of the PRO in the clinical trial be correctly addressed. A checklist of recommendation is provided.

Non-GVHD-Specific PRO Measures

The HCT field has made significant efforts to harmonize the generic PRO measures which should be used in the research setting. The Design of Clinical Trials Working Group [32••] recommended the SF36, FACT-BMT, or HAP measures in adults. The BMT CTN has called for a core set of measurement tools for study endpoints of QOL [43], and more recently, the BMT CTN and the CIBMTR prepared a consensus paper which recommended using a core set of domains from a versatile freely available single measure system, which is easy to access and has a low symptom burden. The NIH PROMIS measures were proposed as the system which met all of these requirements [44]. The group went on to study the PROMIS measures in HCT survivors and compare its usefulness with that of the SF36. Results show a high degree of correlation between the SF36 and the PROMIS scores and suggest that symptoms and function can be adequately assessed using the SF36 or PROMIS measures, both in the general population of HCT survivors [45] and those with GVHD [46].

In aGVHD, the MDASI has been used in some studies [2]. The Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™) [47] is a PRO measure developed to evaluate symptomatic toxicity specifically in the contest of clinical trials. The PROCTCAE is 124 items, which represent 78 symptomatic toxicities, include attributes such as frequency, severity, and interference. The choice of items to include in an individual trial is determined by the expected toxicities that will be experienced and there is no need for a patient to complete the entire battery and no summary score. The measure is freely available. The PROCTCAE may be useful for GVHD clinical trials [35], especially acute GVHD [2], which has a high burden of symptoms which are hard to the clinical team to precisely quantify (e.g., nausea and anorexia) or assess the impact of (interference due to diarrhea). There is limited experience with the PROCTCAE in HCT in general, although some data suggests its use is feasible in early post-HCT even when frequently administered [48]. The PROCTCAE and PROMIS have, not to my knowledge, been reported in GVHD clinical trials and there are no studies listed on clinicaltrials.gov with these search terms and GVHD.

Composite Scales

Composite scales may be of great use in GVHD trials as they capture clinician assessment, PRO, as well as laboratory or functional measures. Composite endpoints have been proposed by investigators in order to fully represent that the clinical benefit obtained in a trial may encompass not only a clearly defined primary endpoint but also benefits in symptoms, functions, and QOL associated with the primary endpoint [31]. However, composite scales are complex to develop, each component needs to be justified and show clinical benefit and sample size

as well as a gold standard against which to validate the scale need to be considered. Currently, no composite scales incorporating PRO in GVHD have been developed.

Conclusions

In summary, a small proportion of clinical trials in GVHD treatment and prophylaxis now include PROs, typically as secondary outcomes, but it is not always possible to find a publication for a completed trial, and publications do not always include the PRO outcomes. Additionally, PRO endpoints are often presented separately to other outcomes and may appear confusing or contradictory and hard to reconcile clinically. Trials are frequently inadequately powered for the PRO outcome (either initially or due to accrual issues) leading to a lack of confidence in the value of the PRO results.

This is despite significant efforts by chronic GVHD investigators to incorporate PROs, and drive a patient-centered approach, in clinical trials for GVHD treatment. Reassuringly, evidence of recent adoption of the consensus recommendations may be seen in recently recruiting trials which now commonly incorporate the GVHD-specific Lee SS and the non-GVHD-specific FACT-BMT and/or SF36. The Lee SS is recommended in therapeutic response trials as a chronic GVHD-specific core measure alongside other clinician and patient assessments and reports. Additionally, non-GVHD-specific measures are strongly recommended as secondary outcomes. Although there is substantial experience with generic measures including the SF36 and FACT-BMT, there are some limitations to their use such as licensing requirements and higher respondent burden. Growing experience with the PROMIS measures in HCT, as well as the benefits the PROMIS system has over the currently used instruments, suggest this may become routinely used in the future.

Less progress has been made in acute GVHD with no published recommendations suggesting a standard approach to PRO collection in trials. A validated PRO measure in this setting is needed and could potentially be built using elements of the PROCTCAE. Pediatrics is also less well represented in the development of specific measures in this area. Finally, widespread support in the field for routine inclusion of PRO as primary (or even secondary) objectives in GVHD trials will require confidence in PRO results which will follow from well-conducted clinical trials performed following consensus recommendations from both GVHD and PRO experts.

Recommendations

1. Incorporate PROs into cGVHD studies according to the recommendations set out by the NIH consensus project for criteria for clinical trials in cGVHD
2. Promote FDA qualification of the Lee Symptom Scale for cGVHD
3. Pursue the development of a validated PRO measure for acute GVHD
4. Follow the CONSORT and SPIRIT PRO extensions to ensure the protocol and publication fully describe the rationale, delivery, and value of the PRO collection
5. Advance both the clinical and statistical development of composite endpoints as primary endpoints in GVHD clinical trials
6. Harmonize the measures used for non-GVHD-specific PRO collection

Compliance with Ethical Standards

Conflict of Interest The author declares no conflict of interests.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Papers of particular interest, published recently, have been highlighted as:
•• Of major importance

1. Brundage M, Bass B, Davidson J, Queenan J, Bezjak A, Ringash J, et al. Patterns of reporting health-related quality of life outcomes in randomized clinical trials: implications for clinicians and quality of life researchers. Qual Life Res. 2011;20(5):653–64.
2. Lee SJ, Williams LA. Patient-reported outcomes for acute graft-versus-host disease prevention and treatment trials. Biol Blood Marrow Transplant. 2010;16(3):295–300.
3. Martin PJ, Weisdorf D, Przepiorka D, Hirschfeld S, Farrell A, Rizzo JD, et al. National Institutes of Health Consensus Development Project On Criteria For Clinical Trials in chronic graft-versus-host disease: VI. Design of Clinical Trials Working Group report. Biol Blood Marrow Transplant. 2006;12(5):491–505.
4. FDA. US Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. US Department of Health and Human Services Food and Drug Administration. 2009. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Accessed 30 Oct 2018.
5. Rothrock NE, Kaiser KA, Cella D. Developing a valid patient-reported outcome measure. Clin Pharmacol Ther. 2011;90(5):737–42.
6. McQuellon RP, Russell GB, Cella DF, Craven BL, Brady M, Bonomi A, et al. Quality of life measurement in bone marrow transplantation: development of the functional assessment of cancer therapy-bone marrow transplant (FACT-BMT) scale. Bone Marrow Transplant. 1997;19(4):357–68.
7. Cleeland CS, Mendoza TR, Wang XS, Chou C, Harle MT, Morrissey M, et al. Assessing symptom distress in cancer patients: the M.D. Anderson symptom inventory. Cancer. 2000;89(7):1634–46.

8. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30(6):473–83.
9. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727–36.
10. Lee S, Cook EF, Soiffer R, Antin JH. Development and validation of a scale to measure symptoms of chronic graft-versus-host disease. Biol Blood Marrow Transplant. 2002;8(8):444–52.
11. Daughton DM, Fix AJ, Kass I, Bell CW, Patil KD. Maximum oxygen consumption and the ADAPT quality-of-life scale. Arch Phys Med Rehabil. 1982;63(12):620–2.
12. Hoodin F, LaLonde L, Errickson J, Votruba K, Kentor R, Gatza E, et al. Cognitive function and quality of life in vorinostat-treated patients after matched unrelated donor myeloablative conditioning hematopoietic cell transplantation. Biol Blood Marrow Transplant. 2018. https://doi.org/10.1016/j.bbmt.2018.09.015.
13. Bolanos-Meade J, Logan BR, Alousi AM, Antin JH, Barowski K, Carter SL, et al. Phase 3 clinical trial of steroids/mycophenolate mofetil vs steroids/placebo as therapy for acute GVHD: BMT CTN 0802. Blood. 2014;124(22):3221–7 quiz 335.
14. Carpenter PA, Logan BR, Lee SJ, Weisdorf DJ, Johnston L, Costa LJ, et al. A phase II/III randomized, multicenter trial of prednisone /sirolimus vs prednisone/sirolimus/calcineurin inhibitor for treatment of chronic graft-versus-host disease: BMT CTN 0801. Haematologica. 2018;103:1915–24.
15. Baird K, Comis LE, Joe GO, Steinberg SM, Hakim FT, Rose JJ, et al. Imatinib mesylate for the treatment of steroid-refractory sclerotic-type cutaneous chronic graft-versus-host disease. Biol Blood Marrow Transplant. 2015;21(6):1083–90.
16. Cutler C, Miklos D, Kim HT, Treister N, Woo SB, Bienfang D, et al. Rituximab for steroid-refractory chronic graft-versus-host disease. Blood. 2006;108(2):756–62.
17. St John L, Gordon SM, Childs R, Marquesen M, Pavletic SZ, Wu TX, et al. Topical thalidomide gel in oral chronic GVHD and role of in situ cytokine expression in monitoring biological activity. Bone Marrow Transplant. 2013;48(4):610–1.
18. DeFilipp Z, Nazarian RM, El-Jawahri A, Li S, Brown J, Del Rio C, et al. Phase 1 study of the hedgehog pathway inhibitor sonidegib for steroid-refractory chronic graft-versus-host disease. Blood Adv. 2017;1(22):1919–22.
19. Pavletic SZ. Response as an end point in treatment trials for acute GVHD. Bone Marrow Transplant. 2012;47(2):161–3.
20. Harris AC, Young R, Devine S, Hogan WJ, Ayuk F, Bunworasate U, et al. International, multicenter standardization of acute graft-versus-host disease clinical data collection: a report from the Mount Sinai Acute gvhd International Consortium. Biol Blood Marrow Transplant. 2016;22(1):4–10.
21. 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 Transplant. 2018;53(11):1401–15.
22. Chronic GC. Rationale and design of the chronic GVHD cohort study: improving outcomes assessment in chronic GVHD. Biol Blood Marrow Transplant. 2011;17(8):1114–20.
23.•• Pidala J, Kurland B, Chai X, Majhail N, Weisdorf DJ, Pavletic S, et al. Patient-reported quality of life is associated with severity of chronic graft-versus-host disease as measured by NIH criteria: report on baseline data from the Chronic GVHD Consortium. Blood. 2011;117(17):4651–7 **This paper shows that cGVHD severity is independently associated with Quality of Life.**
24. Pidala J, Kurland BF, Chai X, Vogelsang G, Weisdorf DJ, Pavletic S, et al. Sensitivity of changes in chronic graft-versus-host disease activity to changes in patient-reported quality of life: results from the chronic graft-versus-host disease consortium. Haematologica. 2011;96(10):1528–35.
25. Inamoto Y, Chai X, Kurland BF, Cutler C, Flowers ME, Palmer JM, et al. Validation of measurement scales in ocular graft-versus-host disease. Ophthalmology. 2012;119(3):487–93.
26. Inamoto Y, Martin PJ, Chai X, Jagasia M, Palmer J, Pidala J, et al. Clinical benefit of response in chronic graft-versus-host disease. Biol Blood Marrow Transplant. 2012;18(10):1517–24.
27. Jacobsohn DA, Rademaker A, Kaup M, Vogelsang GB. Skin response using NIH consensus criteria vs Hopkins scale in a phase II study for steroid-refractory chronic GVHD. Bone Marrow Transplant. 2009;44(12):813–9.
28. Palmer J, Williams K, Inamoto Y, Chai X, Martin PJ, Tomas LS, et al. Pulmonary symptoms measured by the national institutes of health lung score predict overall survival, nonrelapse mortality, and patient-reported outcomes in chronic graft-versus-host disease. Biol Blood Marrow Transplant. 2014;20(3):337–44.
29. Treister N, Chai X, Kurland B, Pavletic S, Weisdorf D, Pidala J, et al. Measurement of oral chronic GVHD: results from the Chronic GVHD consortium. Bone Marrow Transplant. 2013;48(8):1123–8.
30. Merkel EC, Mitchell SA, Lee SJ. Content validity of the Lee chronic graft-versus-host disease symptom scale as assessed by cognitive interviews. Biol Blood Marrow Transplant. 2016;22(4):752–8.
31. 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):1343–59.
32.•• 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):984–99 **This consensus document describes the use of the LEE symptom scale as a chronic-GVHD core measure to use in cGVHD clinical trials.**
33. Pidala J, Chai X, Kurland BF, Inamoto Y, Flowers ME, Palmer J, et al. Analysis of gastrointestinal and hepatic chronic graft-versus-host [corrected] disease manifestations on major outcomes: a chronic graft-versus-host [corrected] disease consortium study. Biol Blood Marrow Transplant. 2013;19(5):784–91.
34. Palmer J, Chai X, Pidala J, Inamoto Y, Martin PJ, Storer B, et al. Predictors of survival, nonrelapse mortality, and failure-free survival in patients treated for chronic graft-versus-host disease. Blood. 2016;127(1):160–6.
35. Efficace F, Gaidano G, Lo-Coco F. Patient-reported outcomes in hematology: is it time to focus more on them in clinical trials and hematology practice? Blood. 2017;130(7):859–66.
36. Efficace F, Fayers P, Pusic A, Cemal Y, Yanagawa J, Jacobs M, et al. Quality of patient-reported outcome reporting across cancer randomized controlled trials according to the CONSORT patient-reported outcome extension: a pooled analysis of 557 trials. Cancer. 2015;121(18):3335–42.
37. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gotzsche PC, Krleza-Jeric K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7.
38. Chan AW, Tetzlaff JM, Gotzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ. 2013;346:e7586.
39. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan AW, King MT, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. JAMA. 2018;319(5):483–94.
40. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c869.

41. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD, et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. JAMA. 2013;309(8):814–22.
42. CONSORT. http://www.consort-statement.org/extensions?ContentWidgetId=560, Accessed 30 Oct 2018.
43. Appelbaum FR, Anasetti C, Antin JH, Atkins H, Davies S, Devine S, et al. Blood and marrow transplant clinical trials network state of the science symposium 2014. Biol Blood Marrow Transplant. 2015;21(2):202–24.
44. Shaw BE, Lee SJ, Horowitz MM, Wood WA, Rizzo JD, Flynn KE. Can we agree on patient-reported outcome measures for assessing hematopoietic cell transplantation patients? A study from the CIBMTR and BMT CTN. Bone Marrow Transplant. 2016;51(9):1173–9.
45. Shaw BE, Syrjala KL, Onstad LE, Chow EJ, Flowers ME, Jim H, et al. PROMIS measures can be used to assess symptoms and function in long-term hematopoietic cell transplantation survivors. Cancer. 2018;124(4):841–9.
46. Lee SJ, Onstad L, Chow EJ, Shaw BE, Jim HSL, Syrjala KL, et al. Patient-reported outcomes and health status associated with chronic graft-versus-host disease. Haematologica. 2018;103(9):1535–41.
47. PROCTCAE. https://healthcaredelivery.cancer.gov/pro-ctcae/. Accessed 30 Oct 2018.
48. Wood WA, Deal AM, Abernethy A, Basch E, Battaglini C, Kim YH, et al. Feasibility of frequent patient-reported outcome surveillance in patients undergoing hematopoietic cell transplantation. Biol Blood Marrow Transplant. 2013;19(3):450–9.