



Published in final edited form as:

Transplant Cell Ther. 2026 June ; 32(6): 644.e1–644.e19. doi:10.1016/j.jtct.2026.01.009.

Towards Better Response Assessment of Cutaneous Chronic Graft-Versus-Host Disease: A Report from the National Institutes of Health Consensus Project Task Force

Alina Markova, MD,

Dermatology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA; Weill Cornell Medicine, New York, NY

Stephanie J. Lee, MD, MPH,

Clinical Research Division, Fred Hutchinson Cancer Center, Seattle, WA; Department of Medicine, University of Washington, Seattle, WA

Badri Modi, MD,

Division of Dermatology, City of Hope, Duarte, CA

Emily Baumrin, MD, MSCE,

Department of Dermatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

Rachel K. Rosenstein, MD, PhD,

Center for Discovery and Innovation and Department of Internal Medicine, Hackensack Meridian School of Medicine, Hackensack University Medical Center, Nutley, NJ

Eric R. Tkaczyk, MD, PhD,

Department of Veterans Affairs, Nashville, TN

Department of Dermatology, Vanderbilt University Medical Center, Nashville, TN

Julia S. Lehman, MD,

Departments of Dermatology and Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN

Yoshihiro Inamoto, MD,

Department of BMT and Cellular Therapy, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

Paul A. Carpenter, BSc MBBS,

Clinical Research Division, Fred Hutchinson Cancer Center, Seattle, WA; Department of Pediatrics, University of Washington, Seattle, WA

Corresponding Author: Alina Markova, MD, Vice Chair, Quality, Department of Medicine, Section Head, General Dermatology and Oncodermatology, Memorial Sloan Kettering Cancer Center, 545 E 73rd St, New York, NY 10021, F 929-321-8102, markovaa@mskcc.org.

*Co-Senior Authors

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Kirk R. Schultz, MD,

Michael Cuccione Childhood Cancer Research Program, British Columbia Children's Hospital Research Institute, University of British Columbia, Vancouver, BC, British Columbia

Joseph A. Pidala, MD, PhD,

Department of Blood and Marrow Transplantation and Cellular Immunotherapy, Moffitt Cancer Center, Tampa, FL

Paul Martin, MD,

Clinical Research Division, Fred Hutchinson Cancer Center, Seattle, WA; Department of Medicine, University of Washington, Seattle, WA

Iskra Pusic, MD,

Division of Oncology, Washington University School of Medicine, Siteman Cancer Center, St. Louis, MO

Sencer Goklemez, MD,

Immune Deficiency Cellular Therapy Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD; Division of Hematology Oncology, Department of Internal Medicine, University of Cincinnati Medical Center, Cincinnati, OH

Christian Stockmann, PhD,

University of Zurich, Institute of Anatomy, Zurich, Switzerland

Iago Pinal-Fernandez, MD, PhD,

Muscle Disease Section, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD. Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD

Sophie Paczesny, MD, PhD,

Hollings Cancer Center and Department of Pharmacology and Immunology, Medical University of South Carolina, Charleston, SC

Andrew Harris, MD,

Pediatric Transplant and Cellular Therapies, Memorial Sloan Kettering Cancer Center, New York, NY

Doris M. Ponce, MD, MS,

Department of Medicine, Memorial Sloan Kettering Cancer Center; New York, NY. Weill Cornell Medical College, New York, NY

Noa G. Holtzman, MD,

Division of Transplantation and Cell Therapy, Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL

Najla El Jurdi, MD,

Immune Deficiency Cellular Therapy Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD

Steven Z. Pavletic, MD, PhD*,

Immune Deficiency Cellular Therapy Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD

Edward W. Cowen, MD, MHSc*

Dermatology Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases,
National Institutes of Health, Bethesda, MD

Abstract

The National Institutes of Health (NIH) chronic graft-versus-host disease (cGVHD) Consensus Project established response criteria that enabled clinical trials and facilitated regulatory approval of multiple therapies. Nonetheless, organ-specific assessments have limitations, particularly for severe sclerotic skin involvement. Cutaneous manifestations occur in approximately half of patients with chronic GVHD but are difficult to evaluate due to heterogeneous clinical features. To address these challenges, the NIH Consensus Skin Task Force convened in 2024–2025 to refine skin response measures for use in clinical trials. This report (a) summarizes current diagnosis and scoring of skin chronic GVHD, (b) reviews existing response assessments, (c) identifies gaps in their performance, (d) proposes refinements to the 2014 NIH skin response criteria, and (e) outlines future directions incorporating patient-reported outcomes, novel technologies, and biomarker research. Current NIH scoring relies on a 4-point body surface area (BSA) scale and a 3-point sclerosis features scale. While standardized, these measures have limited sensitivity to clinically meaningful change. Proposed refinements include: (i) separate BSA assessments for epidermal involvement and sclerotic features; (ii) modification of sclerosis descriptors, with removal of impaired mobility and specification of GVHD-related ulceration, and addition of sclerosis-associated edema with or without erythema; and (iii) replacement of the exploratory severity scale with two new clinician instruments. The Sclerosis Quality and Physical Signs (SQPS) scale (0–10) captures qualitative physical changes, while the Sclerosis Daily Function Impact (SDFI) scale (0–4) evaluates functional compromise. These proposed refinements aim to improve the accuracy, reproducibility, and clinical relevance of skin chronic GVHD assessments, strengthening trial endpoints and patient care.

Keywords

Chronic Graft-Versus-Host Disease; Chronic Cutaneous GVHD; Skin GVHD; Sclerotic GVHD; Epidermal GVHD; Chronic GVHD Response

Introduction

The NIH chronic graft-versus-host disease (cGVHD) Consensus project established the response criteria and framework for clinical trials, which facilitated the eventual regulatory approvals of multiple FDA-approved therapies for chronic GVHD[1–4]. Notwithstanding these advances, improved organ-specific response measures are needed, especially for severe sclerotic skin manifestations. The skin is affected in approximately 50% of patients with chronic GVHD during their disease course, but is difficult to accurately assess, in part due to its highly variable appearance [5, 6]. A NIH Chronic GVHD Consensus Conference task force of Dermatology and Hematopoietic Cell Transplant (HCT) specialists was convened in April 2024 to propose and refine measures to more accurately and reliably assess skin chronic GVHD responses in clinical trials. Multiple revisions were discussed via conference calls. The task force met again in person at the International chronic GVHD Symposium

in Vancouver, BC (November 2024) and a third time for public comments at the American Society of Transplantation and Cellular Therapy Annual Meeting in Honolulu, HI (February 2025).

Purpose of this document

The purpose of this document is to: (a) summarize current diagnosis and scoring of cutaneous chronic GVHD; (b) review current cutaneous chronic GVHD response assessments; (c) define gaps in performance of these assessments in routine care and clinical trials; (d) propose refinements in cutaneous chronic GVHD response tools; and (d) offer clinicians and researchers actionable strategies to further refine the accuracy of these assessments through the integration of patient-reported outcomes, innovative technologies, and biomarker research.

Current skin chronic GVHD diagnosis and scoring

Current skin NIH severity scoring consists of two coarse measurement scales. First, a 4-point “*Score % BSA*” scale categorizes the total body surface area (BSA) involved, which is based on any combination of maculopapular rash/erythema, lichen planus-like, sclerotic features, papulosquamous lesions, or ichthyosis, and keratosis pilaris-like GVHD. A score of 0 indicates no involvement, while a score of 1 indicates 1–18% BSA involvement, a score of 2 equates to 19–50% BSA, and a score of 3 > 50% BSA. The second “*Skin features score*” 3-point scale broadly categorizes the extent of sclerotic skin features based on the presence or absence of superficial or deep sclerosis: a score of 0 is given if no sclerosis is present, whereas superficial sclerosis is given a score of 2, and deep sclerosis, hidebound, impaired mobility, and/or ulceration is given a score of 3. Notably, in recognition that a diagnosis of sclerosis immediately connotes potentially highly morbid chronic GVHD, “score 1” is absent from the skin features scale. When a discrepancy exists between the percentage of total BSA score and the skin feature score, the higher number is used for the final skin score (Figure 1A–B [7]).

Epidermal vs sclerotic GVHD

Significant interobserver variability has been noted in the scoring of cutaneous chronic GVHD because the manifestations can be pleomorphic [8]. Epidermal skin GVHD includes maculopapular rash/erythema, lichen planus-like, papulosquamous lesions, ichthyosis, and keratosis pilaris-like GVHD—that is, manifestations in which skin pathology reveals the presence of GVHD in the epidermis. Erythema is often but not always present and varies significantly by patient skin pigmentation. Lichen planus-like changes are the prototypic epidermal manifestations diagnostic of chronic GVHD and are thought to arise from T-cell-mediated attack of basal layer epidermal keratinocytes. Clinically, localized or generalized pink to violaceous papules or plaques with overlying scale are seen. Erythema and/or maculopapular rash alone without tissue changes suggestive of epidermal GVHD should not be scored in this category in the absence of confirmatory skin biopsy.

In contrast, sclerotic chronic GVHD refers to superficial or deep sclerosis, representing all chronic GVHD skin manifestations with abnormal dermal or subcutaneous fibrosis. Lichen sclerosus-like chronic GVHD involves inflammation and sclerosis in the superficial

dermis, leading to gray or ivory-white thin plaques, often involving the anogenital region, neck, and back. Sclerosis of the deeper dermis and fascia may result in localized plaques with or without erythema resembling morphea, diffuse skin thickening resembling systemic sclerosis (SSc), or a rippled and nodular appearance resembling subcutaneous morphea or eosinophilic fasciitis. Sclerotic skin manifestations of chronic GVHD may be preceded by soft tissue edema due to the inflammatory phase of sclerosis, though edema unrelated to chronic GVHD in the setting of concurrent medications, venous insufficiency or other causes must be considered, further complicating accurate skin assessment. As erythema may accompany both epidermal and sclerotic GVHD, may be difficult to detect in darkly pigmented skin, and is not in itself a diagnostic chronic GVHD manifestation, we recommend avoiding the term ‘erythema’ and referring to the skin cGVHD subtypes simply as *epidermal* or *sclerotic*. In addition, if superficial skin changes are attributed exclusively to underlying sclerosis (e.g., scaling skin or wrinkled skin changes overlying lichen sclerosus/morphea) without features suggestive of active epidermal GVHD, then these areas should be calculated as sclerotic involvement only.

Determination of ongoing GVHD skin activity vs damage

For clinical practice, it is also desirable to differentiate skin features that represent *active* ongoing immune-mediated tissue injury from skin manifestations that reflect past tissue *damage* (e.g., sequelae of prior inflammatory disease) as activity drives management of systemic therapy. Transition of active skin chronic GVHD to a quiescent state often results in residual skin changes, which may be slow to resolve or result in permanent skin changes[9]. These changes may involve the skin (post-inflammatory hyper- or hypopigmentation, scarring at previous ulcer site), hair (scarring or non-scarring alopecia), or nails (nail dystrophy or loss). These skin changes are not scored for the purposes of BSA skin response assessment despite being well-established sequelae of cutaneous chronic GVHD. Currently, no standard clinical scales or biomarker assays exist to differentiate “active” versus “damage” in chronic GVHD, which represents an important area of unmet need for research.

Utility of skin biopsy

Skin biopsy is a low-morbidity diagnostic tool for histologic confirmation of suspected chronic GVHD and is supported by a prior multidisciplinary consensus statement[10] and the National Comprehensive Cancer Network clinical practice guidelines for chronic GVHD[11]. In patients with suspected chronic GVHD, skin biopsy examination may offer value for: (1) diagnostic confirmation of GVHD in the absence of clinically diagnostic features and exclusion of other entities such as malignant/pre-malignant lesions, drug reaction, or infection, (2) disease subtype classification when clinical findings are non-diagnostic, and (3) assessment of extent of ongoing disease activity/prognosis[12]. GVHD pathology should be reported according to the NIH 2014 Pathology Working Group Report[13].

First, the presence of certain histopathologic features can support a diagnosis of chronic GVHD and correlate with disease subtypes. Specifically, vacuolar interface alteration (characterized by the presence of cytoid bodies and hydropic degeneration of the basal

layer) along the dermal-epidermal junction and extending down hair follicles correlates with active epidermal disease[10, 13, 14]. Even if vacuolar interface changes are not present, the presence of melanin incontinence (superficial dermal pigmented melanophages) is suggestive of prior interface changes at the dermal-epidermal junction and may be the sole histopathologic finding in chronic GVHD-associated hyperpigmentation[10]. Sclerotic lesions demonstrate various degrees of dermal, pannicular septal, or fascial sclerosis, along with aggregates of lymphoplasmacytic inflammation[13].

Though histopathologic grading is not currently incorporated in chronic GVHD clinical scoring[13], the skin represents an accessible tissue with potential as a prognostic tool. While rigorous studies correlating skin histopathologic features with clinical outcomes have been performed for morphea[15] and systemic sclerosis[16], data evaluating the role of skin biopsy for chronic GVHD disease activity are relatively sparse[13]. One group found no correlation between depth of sclerosis and patient outcomes with chronic GVHD; however, this study did not consider other histopathologic variables such as the presence of dermal/epidermal interface changes or degree of inflammation[17].

Measuring Response in Chronic Cutaneous GVHD

Current Response Criteria: In clinical trials, individual target organ responses to treatment have been classified using the following three categories from the 2014 NIH Consensus Criteria into: complete organ response, partial organ response, and lack of response (includes the subcategories of unchanged and organ progression)[7]. For skin, a *complete response* is defined as a change in organ response score from ≥ 1 to 0, *partial response* is defined as decrease in score by at least one point but with persistent disease (score >0), *unchanged* as no change from previous score, and *progression* as an increase in score by at least one (except 0 to 1).

The original 2005 NIH Response Criteria proposed to assess skin involvement by scoring three discrete BSA categories (erythematous rash, movable sclerosis, nonmoveable sclerosis). However, BSA measurement suffers from poor inter-rater reliability, particularly for sclerotic skin features[8]. The 2014 Consensus Conference revision acknowledged these limitations and introduced a 0–3 scale based on broad ranges of BSA involvement (Figure 1A)[7]. However, this methodology is significantly limited when arbitrary category boundary definitions interact with a naturally continuous variable like BSA involvement, as demonstrated in Table 1 [8]. To improve the sensitivity to change in skin sclerosis, an exploratory clinician-reported global assessment of skin and/or joint tightening severity was also proposed in 2014 (0 indicated no GVHD manifestations, 10 indicated the most severe chronic GVHD symptoms possible)[7]. Although this scale might be useful to detect clinically meaningful change in sclerosis that is not captured by BSA, it is not currently used as the standard for the determination of overall skin response.

Current NIH skin response assessment may, therefore, potentially fail to detect large changes in BSA, or conversely, falsely lead to re-classification of small, reported changes in BSA as either response or disease progression. Importantly, current methodology does not ensure that a change in skin organ score reflects a clinically relevant change for the patient. For instance, subjective skin softening may provide meaningful clinical benefit to the patient

(decreased pain, tightness, and improved flexibility) yet may not be reflected by a detectable change in sclerosis in the skin features score. Therefore the 2014 NIH chronic GVHD Consensus recommended that trials targeting sclerotic chronic GVHD use “more detailed measures to document change in extent and functional consequences of sclerosis”[7]. Additionally, the 2020 NIH Consensus Committee report on highly morbid forms of chronic GVHD noted that a lack of worsening, improved functionality, and improvement in patient-reported outcomes represent clinically meaningful benefit for a given patient even if chronic GVHD does not improve by objective measures[18].

Lastly, accumulating data highlights the prognostic impact of accurate capture of cutaneous disease subtypes and the extent of involvement. The BSA of epidermal disease is predictive of poorer survival, suggesting that accurate BSA measurement is important for prognosis [19–21]. It is conceivable that the difficulty with assessing “active” versus “inactive” sclerotic disease at any given time point may also contribute to the poor performance of sclerosis as a reliable prognostic finding. In summary, a more accurate assessment of epidermal and sclerotic skin manifestations at diagnosis and throughout the disease course would allow for improved response determination and prognostication.

Summary of current NIH skin assessment performance in clinical trials

Four therapeutic agents – ibrutinib, belumosudil, ruxolitinib, and axatilimab – have received FDA approval to date for the treatment of chronic GVHD based on overall response rates observed in clinical trials. Apart from the ibrutinib trial, in which response assessments were based on the modified 2005 NIH consensus criteria, each clinical trial utilized Form A (Clinician Report) from the 2014 NIH consensus criteria to assess cutaneous chronic GVHD response. Given that these studies examined overall chronic GVHD response rate (ORR), improvement in any organ allowed for partial response scoring, with or without skin involvement at baseline or associated cutaneous improvement. Thus, overall response rates (ORRs) in clinical studies tended to be higher than skin-specific response rates (Table 2). The current reporting guidelines do not allow conclusions to be drawn about response rates of chronic GVHD cutaneous chronic GVHD subtypes, notably epidermal and superficial or deep sclerosis. Apart from the axatilimab trial, the other three major clinical trials report a composite skin response without specifying baseline skin phenotypes (presence, absence, change in features) and the specific metric used to determine response or no-response. The axatilimab study specifically reported that 44.2% of patients with non-moveable sclerosis had any improvement in their % BSA.

Considerations for clinical trials of topical or skin-directed therapy

Assessment of topical or other skin-directed treatment efficacy for cutaneous chronic GVHD is often complicated by the effect of concurrent systemic therapy. Skin-directed therapeutic trials may benefit from designs such as the use of a split-body application approach to reduce confounding from concurrent systemic therapy[22–24]. Several prospective topical interventions have been evaluated for the treatment of cutaneous chronic GVHD. A 2001 study of split-body tacrolimus ointment defined response as effective subjective relief of erythema and/or pruritus[25]. A second 2003 tacrolimus ointment study of 10 patients subjectively graded response 0–4 (from no response [no change in GVHD signs], with

intermediate responses based on percent reduction of involved surface area, to complete response [of GVHD][26]. Finally, a randomized controlled trial of topical ruxolitinib utilized three objective scoring tools-BSA, physician global assessment (PGA), and composite assessment of index lesion severity (CAILS), and detected statistically significant improvements with all scoring tools[22, 23]. The PGA showed a correlation with other more detailed response measures, such as BSA reduction and CAILS in the ruxolitinib cream study, demonstrating that physicians' assessment on a 0–6 scale is a promising rapid assessment tool for cutaneous chronic GVHD response.

Challenges in the diagnosis and assessment of pediatric cutaneous chronic GVHD

Limitations in the application of NIH consensus criteria for the diagnosis of chronic GVHD in children have been described[27]. Although cutaneous manifestation may resemble those in adults, quantifying skin involvement is more challenging due to evolving BSA distributing during growth (Supplement, Figure 2). While some tools estimate pediatric BSA[28], none account for its dynamic changes, limiting accurate scoring and response assessment compared with adults. Growth abnormalities seen in juvenile scleroderma [29] raise questions about whether sclerotic GVHD similarly affects development. Skin biopsy use is limited in children due to procedural concerns, and histologic findings remain largely extrapolated from adults. Moreover, while several PRO tools are validated in adults, none exist for pediatric GVHD, hindering evaluation of quality of life and treatment response in this population. Finally, age-related differences in skin must be considered when validating novel assessment tools such as the Durometer and Myoton.

Methods for estimation of body surface area (BSA)

The clinical estimation of BSA of skin diseases has been derived primarily from several historical methods (Supplement Table 1)[30]. The contributing BSA of anatomic regions in these standard approaches is based on paper mâché models of 12 humans from the early 20th century, which also produced the widely-used DuBois formula for BSA[31]. More recent data, derived from laser scanning more than 3000 adults in the Civilian American and European Surface Anthropometry Resource, demonstrated that human proportions vary much more widely than previously assumed[32]. All methods suffer from major errors and inter-observer variation. Generally, the more training, time, and effort a BSA estimation protocol requires, the more reliable the assessment will be.

In the future, emerging imaging and digital approaches, including SmartPhone apps, 3D imaging, computer vision approaches, and/or increasingly refined digital avatars, may offer a more efficient approach to overcome these difficulties. Of 31 smartphone apps intended for use by patients with burns systematically reviewed in 2015[33], 19 estimated BSA, with many subsequent advances in ease of use and accuracy compared to clinical methods[32, 34, 35]. Finally, given ongoing advances in machine learning[36], direct application of artificial intelligence image analysis to chronic GVHD clinical images[37] may allow for automated estimation of chronic GVHD skin involvement.

While this task force does not suggest one specific method to calculate BSA, we recommend using the most precise and reliable BSA estimation method that can be practically supported

by the study design. The selected methodology should be clearly defined in advance in the study protocol and uniformly applied by all study assessors. Amongst several sources of error in BSA estimation[38], human calculation error is a significant contributor. Accuracy is enhanced when assessors input raw data for automated analyses, avoiding manual calculations. Ideally, marked avatars would be collected along with data on the raters' experience and training for BSA estimation, with additional collection of standardized skin photographs. Such photo documentation could also be re-analyzed in the future with emerging image analysis technologies.

Proposed Refinements of the 2014 NIH Skin Response Criteria

To address the limitations of the 2014 NIH skin chronic GVHD response criteria, the task force has proposed a revised response assessment form (Figure 2A).

Response Scales

A. BSA Determination:

1. Separate the assessment of epidermal and sclerotic cutaneous chronic GVHD, including the separate BSA calculation for: (i) epidermal changes and (ii) for sclerotic features (including superficial and deep sclerosis combined).
2. While a BSA calculator tool is provided (Figure 2A), clinicians are encouraged to use the BSA calculation tool of their preference (Supplement Table 1). A pediatric formula is also provided in the supplement.

B. Skin Features assignments:

1. No changes are suggested for the epidermal skin features.
2. For sclerotic skin features, it is proposed:
 - i. Remove 'impaired mobility' due to its overlap with joint chronic GVHD.
 - ii. Modify 'ulceration' to 'GVHD-related ulceration' due to the multifactorial etiology of ulcers.
 - iii. Add 'sclerosis-associated edema +/- erythema' to capture sclerotic erythema in chronic GVHD.

C. To measure sclerosis severity and clinically relevant qualitative changes, two new clinician scales are now proposed in lieu of the prior exploratory 0–10 severity scale:

1. "Sclerosis Quality and Physical Signs" (SQPS) scale:

The 0–10 point SQPS scale will assess clinician-assessed qualitative changes, including, but not limited to: skin softness/pinchability, lymphedema, hair growth, skin tightness, abdominal fullness, dimpling/rippling, etc.

2. “Sclerosis Daily Function Impact” (SDFI) scale:

The 0–4-point SDFI scale will assess the impact of sclerotic GVHD on daily function.

The prior exploratory scale combined the clinician’s assessment of functional and symptom severity, including evaluation of joint tightening. To more precisely distinguish between the clinician’s overall assessment of sclerosis severity and its specific impact on the patient’s daily functioning, the exploratory scale was subdivided into two distinct scales. Clinician assessments of skin chronic GVHD response have been shown to correlate closely with validated response scales[49]. Each scale is proposed to be incorporated into the overall response assessment to test them in clinical practice. Given the pleomorphic nature of sclerotic manifestations, the establishment of strict objective guidance is challenging; consequently, these scales rely on clinician judgment.

- D.** Response assessments will now be separately incorporated into an epidermal cutaneous chronic GVHD response and a sclerotic cutaneous chronic GVHD response as detailed in Figure 2B.

Response Determination

We propose modifying the cutaneous chronic GVHD overall response determination to calculate individual responses for epidermal GVHD and sclerotic GVHD.

A minimum 20% relative change from baseline skin GVHD involvement was selected, as prior multi-center pilot work estimated that the threshold for exceeding measurement error varies by phenotype and is approximately 20%[8]. In addition, an absolute 10% change in BSA is required to ensure that observed changes are clinically meaningful. Group consensus determined that a 10% absolute change reflects a clinically significant difference, whereas relative changes that meet the 20% threshold but correspond to <10% absolute difference from baseline are not considered clinically relevant. Furthermore, in a large multi-center chronic GVHD cohort, each 10% increase in epidermal GVHD BSA was independently associated with worse outcomes, supporting a 10% absolute change as a clinically meaningful difference[19].

Consensus Task Force Recommendations

With the goal of improving response assessment reporting, we advise more detailed reporting of epidermal and sclerotic skin assessment in future clinical trial publications. Preferably, clinical trials should report at least the following data at BOTH baseline and response assessment at a minimum:

- Skin feature score
- BSA of epidermal chronic GVHD skin
- BSA of sclerotic chronic GVHD skin
- Sclerosis Quality and Physical Signs (SQPS) scale
- Sclerosis Daily Function Impact (SDFI) scale

- PROs: The Lee Symptom Scale
- P-ROM (photographic range of motion) scores
- Full body clinical photographs (optional)

Innovative and Emerging Response Tools

Patient-reported outcome measures for skin chronic GVHD: Patient-reported outcomes (PROs) are valuable tools for assessing treatment response from the patient's perspective. Chronic skin GVHD is associated with high symptom burden [50], [51, 52], impaired health-related quality of life [50], and increased progression to long-term disability [53]. PROs predict long-term outcomes in skin chronic GVHD with greater symptom burden and worse health-related quality of life associated with a higher risk of nonrelapse mortality independent of clinical severity [50] [51]. Although correlated, PROs reflect disease constructs that differ from clinician-reported outcome measures and provide complementary information. In steroid-refractory chronic GVHD, belumosudil produced higher response rates in the Lee Symptom Scale skin subscale (70%) than in the NIH Skin Score (36%), highlighting the greater sensitivity of PROs as adjuncts to objective measures of skin disease severity[54]. The 2014 NIH consensus criteria recommended core and ancillary PROs for chronic GVHD trials, but FDA guidance requires validation in the target population and specific trial design [7] [55]. As a result, PROs from general or multisystem chronic GVHD studies may not optimally capture skin-specific disease or its subtypes, and no validated PRO tools exist for children or caregivers, leaving a major gap in understanding pediatric chronic GVHD [50, 56, 57].

Lee Symptom Scale: The Lee Symptom Scale (LSS) is a validated tool for assessing patient-reported symptoms and change in chronic GVHD. It evaluates skin discoloration, rashes, thickened skin, skin ulceration/sores, pruritus, and joint limitations on a 5-point Likert scale and is the current standard PRO in chronic GVHD clinical trials.

The LSS skin subscale correlates with clinician-reported NIH skin score [54] and the LSS muscle/joint subscales correlating with patient-reported changes in sclerotic skin [58] and joint/fascia involvement [59], demonstrating good responsiveness to change. However, qualitative studies suggest that these subscales may not capture the spectrum of symptoms, functional impairments, and quality of life impacts in epidermal and sclerotic disease, raising concerns with content validity for skin involvement [60]. While the LSS should remain standard in systemic and skin chronic GVHD trials, additional items are needed to fully capture the patient experience and detect therapeutic benefit.

Other patient-reported outcome measures: Other validated tools, such as Medical Outcomes Study Short Form 36 (SF36), and Patient-Reported Outcomes Measurement Information System (PROMIS) Global and PROMIS-29 scales[61–63] query limitations in activities which may or may not be attributable to skin involvement (e.g., carrying groceries, climbing stairs, etc.) and have not been responsive to change in chronic GVHD studies. Dermatology-specific PROs have not been studied or validated in patients with skin chronic GVHD [64]. The Skindex and Dermatology Life Quality Index are two widely

used measures of health-related quality of life for skin disorders[65, 66]. Although validated in acne and psoriasis, these PROs lack content validity in autoimmune connective tissue diseases, which closely resemble skin chronic GVHD, highlighting the need for disease-specific tools[67, 68]. Unfortunately, despite the clinical similarity between sclerotic chronic GVHD and systemic sclerosis, the Scleroderma Health Assessment Questionnaire was found not to be responsive to clinical changes in a clinical trial of imatinib and rituximab for sclerotic chronic GVHD[69]. Therefore, existing dermatology-specific and generic PROs are not recommended for chronic GVHD clinical trials without further validation.

Skin chronic GVHD PRO development: Efforts in PRO development should focus on disease-specific measures to overcome the aforementioned limitations of existing PROs with regard to content validity and responsiveness to change. A sclerotic chronic GVHD-specific symptom PRO, the Lee Symptom Scale-Skin Sclerosis, has been developed[70]. After psychometric validation, the Lee Symptom Scale-Skin Sclerosis will fill an unmet need to comprehensively capture patient-reported impact of sclerotic chronic GVHD in clinical trials. A skin chronic GVHD-specific PRO addressing symptoms, functional, social, and emotional impacts is in development [71]. These two PROs will be critical adjuncts to the Lee Symptom Scale for chronic GVHD clinical trials [70]. Composite scales integrate clinician, patient, and functional measures, capturing broad benefits in disease activity, but may be less sensitive to changes in individual domains [72, 73]. Future studies should evaluate the role of PROs as co-primary, secondary, or composite endpoints in skin chronic GVHD[74].

Exploratory Tools for Measurement of Sclerotic Chronic Cutaneous GVHD: No validated instruments exist for the objective assessment of sclerotic chronic GVHD. The 2020 NIH Consensus urged adapting tools from other fibrosing diseases, such as the modified Rodnan skin score from SSc, though its role in chronic GVHD requires study [18]. The modified Rodnan skin score has been used to quantify the extent and severity of sclerotic involvement and to support response assessment in several clinical trials in SSc, but its use requires further exploration in sclerotic chronic GVHD[75]. Imaging approaches (high-frequency ultrasound, acoustic radiation force impulse and shear wave elasticity, optical coherence tomography/elastography, MRI with or without PET imaging, OCT laser doppler flowmetry, confocal microscopy, and 3-dimensional photography with computer-assisted quantification), laser doppler flowmetry, confocal microscopy, and 3-dimensional photography with computer-assisted quantification[76–80]. Additional devices that have been reported in the GVHD literature include the Durometer and suction probes such as the Cutometer, Dermaflex, and Nimble[79, 81]. Most modalities have only been studied in small series; developing and validating a rapid, safe, reliable, and accessible objective measure, usable in both topical and systemic trials with minimal training, remains a priority.

Myotonometry: One promising emerging method is tonometry with the Myoton device, which uses a small mechanical impulse to measure multiple biomechanical parameters in under 1 second. The Myoton has had a series of small-scale validations in both SSc[81, 82] and chronic GVHD, demonstrating reproducibility[81, 83], ability to distinguish sclerotic chronic GVHD patients from controls [84, 85], and quantification of changes in chronic

GVHD over time that were largely consistent with the clinician impression of patient response[86].

MRI: MRI may be a useful adjunct to evaluate soft-tissue fibrosis and suspected muscle and fascial abnormalities before the development of extensive range-of-motion limitations[87]. MRI with contrast performed at 1.5 T, 3 T, including T1-weighted spin echo, STIR, and/or 3-dimensional gradient detected clinically detectable sclerosis in chronic GVHD. However, MRI does not consistently detect clinically appreciated changes in cutaneous sclerosis[80, 88]. Therefore, MRI may lack the sensitivity required to effectively serve as a tool for monitoring cutaneous sclerosis treatment response.

Skin tape stripping: Adhesive skin tape stripping (STS) is a minimally invasive technique that allows for the collection of stratum corneum tissue and genomic material from epidermal keratinocytes, T cells, dendritic cells, and melanocytes[23, 89]. In immune-related cutaneous adverse events secondary to immune checkpoint inhibitors, STS correlated with blood biomarkers and skin tissue biomarkers and was successfully used to characterize lichenoid, vitiligo, maculopapular, and psoriasiform subtypes[90]. STS can detect molecular expression profiles correlating with response to topical anti-GVHD JAK inhibition [22, 23]. Notably, STS has effectively detected a pre-inflammatory state in nonlesional psoriatic skin, not captured by skin biopsy, and may have utility in detecting subclinical dermatologic disorders [91]. Noninvasive STS may be explored for early detection of cutaneous GVHD, for monitoring of response, and/or replacing skin biopsy.

Tensile strength management/dye for ECM remodeling: Noninvasive methods to measure skin tensile strength, developed in scleroderma, reflect collagen remodeling and disease stage, and may enable monitoring of sclerotic skin GVHD [92–94]. Advances in probe design allow visualization of extracellular matrix (ECM) remodeling or strain on biopsy material. Fluorescent collagen peptide sensors detect lysyl oxidase activity, localizing active collagen formation beyond conventional immunohistology [95]. Fibrosis is also marked by fibronectin accumulation and altered mechanics. A nanoprobe using the bacterial fibronectin-binding peptide FnBPA5 visualizes fibronectin strain, while radiolabeled ¹¹¹In-FnBPA5 has enabled noninvasive monitoring in murine models [96]. Such probes should be investigated as biomarkers and response tools to improve the diagnosis and monitoring of sclerotic chronic GVHD.

Noninvasive biopsy for fasciitis/myositis: Beyond conventional biopsy, transcriptomics and high-plex proteomics enable sensitive, biologically relevant measurements from smaller samples, allowing serial analysis of transcripts or metabolites. These techniques also permit simultaneous assessment of GVHD-affected sites, including epidermis, dermis, fascia, and muscle. Such advanced methods have driven major insights into systemic autoimmune diseases [97] and immune-related adverse events from immune checkpoint inhibitors[98], and hold promise for application in chronic GVHD assessment to enhance mechanistic understanding and biomarker development..

Skin Biopsy: Future Directions: While exploratory, appraisal of certain microscopic features on skin biopsy examination has the potential to be used to individualize an

approach to therapy. Multiple and complex arms of immunity (i.e., T helper (Th)1, Th2, Th17, Treg, B[103, 104] and macrophage profiles[105]) contribute to chronic GVHD pathogenesis to varying extents, and in many cases, the dominant inflammatory milieu can often be determined from certain microscopic or immunophenotypic findings. For example, vacuolar interface inflammation, present usually in acute and epidermal chronic GVHD, cutaneous lupus, and lichen planus, provides circumstantial evidence of type I interferon-mediated T-cell immunity [106–110]. This inflammatory pattern generally correlates with diseases that respond to JAK inhibition, for example[111]. Similarly, interleukin-36 γ immunohistochemistry has emerged as a potential tissue-based test to assess for Th17 immunity in skin (as is seen in psoriasis, for example) [109, 110] but has not been specifically tested in chronic GVHD. Finally, prognostic tissue-based molecular biomarkers have been explored[112]. Given the increasingly available and affordable high-throughput molecular technology, tissue-based molecular analyses should become more feasible in the clinical setting and are a prime area for innovation and translation to clinical care.

Biomarker development for skin chronic GVHD: Many biomarkers have been identified in patients with cutaneous chronic GVHD, but none are in clinical use. Before adoption in trials, candidates must be validated in both discovery and independent verification cohorts[113]. The chronic GVHD biomarker task force recommends National Institutes of Health–Food and Drug Administration BEST terminology [114]; here, we classify biomarkers as diagnostic, risk/susceptibility, prognostic, predictive, or response/pharmacodynamic.

We propose three main approaches. First, serum/plasma should be collected in trials with detailed clinical data on skin involvement and response. Such biospecimens could support predictive and pharmacodynamic biomarker studies that correlate serological parameters with skin histopathology, including inflammation and fibrosis scores [115] [116]. Collaborative agreements, such as those used in a recent belumosudil study for oral GVHD[117], could facilitate shared use by academic and industry investigators. With expanding therapies, predictive and pharmacodynamic biomarkers will be essential for selecting treatment according to disease stage.

Second, multi-center prospective studies should collect PBMCs and plasma/sera at defined time points—before disease onset, at diagnosis, prior to therapy, post-therapy response, and late disease—together with metadata. Skin biopsies should be collected before therapy and after response, and when feasible, before disease onset for discovery. These samples, banked across institutions, would enable validation of candidate biomarkers and discovery through spatial transcriptomics/proteomics and blood proteomics. Artificial intelligence and machine learning will likely enhance future biomarker development but require large datasets [118]. Finally, validated biomarker candidates should be incorporated into trials to test associations with skin GVHD activity and patient outcomes [119], [120].

Supplement Table 2 highlights biomarkers studied in skin chronic GVHD or fibrosis with the strongest evidence. Candidates limited to murine models or systemic disease without comment on skin involvement were excluded. Most current candidates are diagnostic or risk biomarkers; few have been tested as response biomarkers, representing a critical gap. Two

proteomic studies identified DKK3, a WNT pathway mediator in fibrosis, as a diagnostic [121] and risk/susceptibility biomarker[122] for chronic GVHD with cutaneous sclerosis [123]. Other proteomic work identified MMP3, which degrades extracellular matrix, and CXCL9–10, IFN-inducible chemokines, as risk biomarkers[122]. BAFF has been proposed as a diagnostic biomarker and as predictive of extracorporeal photopheresis (ECP) response [124, 125]. Increased red cell distribution width has also been associated with skin GVHD [126]. The next step is to test these risk biomarkers in trials as activity, predictive, or prognostic biomarkers.

Several cellular candidates are emerging. Blood T cells expressing CXCR3, the receptor for CXCL9–11, are decreased in patients with epidermal or sclerotic chronic GVHD[127]. Regulatory T cells (Tregs), essential for tolerance post-transplantation, are reduced in chronic GVHD but increased in affected skin in proportion to CD8+ T cells [128] [129]. A recent study identified IL-17 and CSF-1R gene signatures in monocytes at or before chronic GVHD diagnosis, which may serve as risk or predictive biomarkers[130].

Skin-based fibrosis biomarkers may offer disease-specific insights and could be translated into soluble or cellular markers. SFRP4, a Wnt regulator, is induced in scleroderma and upregulated in sclerotic chronic GVHD skin and plasma compared to epidermal disease [131–133]. SERPINE2, a TGF- β -regulated serine protease inhibitor [134], and COMP, a PI3K–Akt pathway member, are upregulated in skin and plasma of sclerotic chronic GVHD patients [133]. These candidates should be evaluated in multicenter studies for their potential to track fibrosis activity over time and serve as soluble markers of disease activity.

Inflammatory markers also provide insight into activity. The type I IFN-induced protein MxA is expressed in epidermal and sclerotic chronic GVHD skin[127]. Elafin has been identified as a prognostic biomarker of epidermal chronic GVHD[135]. OX40 ligand is upregulated in sclerotic disease[136], while TREM1 is more prominent in epidermal chronic GVHD[137]. Mast cells and IL-17/IFN- γ -producing CD8+ T cells are enriched in sclerotic and epidermal forms, respectively[103]. CD163+ macrophages contribute to remodeling in acute GVHD, whereas CCR7+ macrophages predominate in chronic disease[105], suggesting CCR7+ cells as risk biomarkers. Prospective studies are needed to determine how such biomarkers might guide therapy escalation or tapering.

Despite substantial progress, no biomarker has yet reached clinical implementation in cutaneous chronic GVHD. Coordinated multicenter studies, harmonized sample collection, and advanced analytic approaches are essential for validation. Integration of fibrosis and inflammatory biomarkers, coupled with AI-driven data analysis, may yield predictive and response biomarkers critical for precision therapy. Ultimately, incorporation of validated biomarkers into clinical trials will be key to improving diagnosis, monitoring, and outcomes in patients with cutaneous chronic GVHD.

CONCLUSIONS

Together, the proposed refinements are designed to improve the sensitivity, reproducibility, and clinical relevance of skin chronic GVHD assessments. By aligning response measures

more closely with patient symptoms and functional outcomes, this framework strengthens clinical trial endpoints, supports regulatory evaluation, and advances patient-centered care. Assessment of the state of research on clinician measures, PROs, adjunctive assessments, and biomarkers provide specific elements to incorporate in novel study designs to get us to the next phase. These goals should be achieved for the field within the next 3–5 years.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS:

This research was supported in part by the Intramural Research Program of the National Institutes of Health (NIH). The contributions of the NIH author(s) were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered Works of the United States Government. However, the findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services. This work was also partially supported by Career Development Award IK2 CX001785 and VA MERIT award I01 CX002721 from the United States Department of Veterans Affairs Clinical Science R&D Service and R01 HL169944 from the National Institutes of Health (R01 HL169944; PI: E. Tkaczyk), National Heart, Lung, and Blood Institute (NHLBI) (R01 HL164902-02; PI: D.M. Ponce), National Cancer Institute (CA118953, CA163438, CA236229; PI: SJ Lee) National Institute for Arthritis and Musculoskeletal and Skin Diseases (K23AR085191; PI: E. Baumrin), NIH K12 Paul Calabresi Award (K12CA226330; N. Holtzman), The Dermatology Foundation Patient Directed Investigation Grant (EB), NIH/NCI Cancer Center Support Grant P30-CA008748, and by SKINTEGRITY.CH consortium, Comprehensive Cancer Center Zurich, Zurich Kidney Center (ZKC), Center for Applied Biotechnology and Molecular Medicine (CABMM) of the University of Zurich (C. Stockmann).

CONFLICTS OF INTEREST:

NONE: Baumrin, Cowen, Holtzman, Rosenstein, Lehman, Pinal-Fernandez, Goklemmez, El Jurdi

Markova: Research Funding: Amryt Pharma, Incyte Corporation, Janssen, Kintara Therapeutics, NIH U01, NIH U54, NIH/NCI Cancer Center Support Grant P30-CA008748, Novartis, Novocure; Consulting: ADC Therapeutics, Alira Health, AstraZeneca, Blueprint Medicines, Incyte Corp, Janssen, OnQuality, Protagonist Therapeutics, Sanofi, Treeline Biosciences; Royalties: UpToDate

Lee: Research Funding: AstraZeneca, Pfizer, Sanofi and Syndax. Consulting: Novartis, Sanofi, Incyte, Orca Bio and Sonoma; and drug supply from Janssen. Clinical trial steering committees for Incyte and Sanofi. Board of Directors of the National Marrow Donor Program (uncompensated)

Paczesny: Holds a patent on “Biomarkers and assays to detect chronic graft versus host disease” (U.S. Patent #10,571,478 B2) that has been licensed to Eurofins/Viracor.

Martin: Equity: Moderna; Procter & Gamble; Honoraria: Pfizer (DSMB); Pro bono clinical trial endpoint adjudication: OrcaBio

Stockmann: Research Funding: Swiss National Fund, Swiss Cancer Research Foundation, The LOOP Zurich, Theiler Haag Foundation, Novartis.

Harris: Consulting/Honoraria: Incyte, Sanofi, Mesoblast, yMabs Therapeutics Inamoto: Research Funding: Meiji Seika Pharma, Japan Society for the Promotion of Science Grant-in-Aid (C) 22K08517; Honoraria: Novartis, Janssen, Meiji Seika Pharma

Ponce: Advisory board member: Incyte, Sanofi, Seres Therapeutics, and Evive Biotechnology Ltd Corporation. Research funding: Incyte and Sanofi.

Pavletic: Research funding: Kadmon, now Sanofi, Eli Lilly, Pharmacyclics, Sobi, Mereo. Consulting/Honoraria: Sanofi, Therakos.

Cowen: Johnson & Johnson Innovative Medicine (employee)

References

1. Wolff D, et al. , Axatilimab in Recurrent or Refractory Chronic Graft-versus-Host Disease. *N Engl J Med*, 2024. 391(11): p. 1002–1014. [PubMed: 39292927]
2. Cutler C, et al. , Belumosudil for chronic graft-versus-host disease after 2 or more prior lines of therapy: the ROCKstar Study. *Blood*, 2021. 138(22): p. 2278–2289. [PubMed: 34265047]
3. Zeiser R, et al. , Ruxolitinib for Glucocorticoid-Refractory Chronic Graft-versus-Host Disease. *N Engl J Med*, 2021. 385(3): p. 228–238. [PubMed: 34260836]
4. Miklos D, et al. , Ibrutinib for chronic graft-versus-host disease after failure of prior therapy. *Blood*, 2017. 130(21): p. 2243–2250. [PubMed: 28924018]
5. Filipovich AH, et al. , National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. *Biol Blood Marrow Transplant*, 2005. 11(12): p. 945–56. [PubMed: 16338616]
6. Jagasia MH, 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): p. 389–401 e1. [PubMed: 25529383]
7. Lee SJ, 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): p. 984–99. [PubMed: 25796139]
8. Mitchell SA, et al. , A multicenter pilot evaluation of the National Institutes of Health chronic graft-versus-host disease (cGVHD) therapeutic response measures: feasibility, interrater reliability, and minimum detectable change. *Biol Blood Marrow Transplant*, 2011. 17(11): p. 1619–29. [PubMed: 21536143]
9. Bhatt NS, et al. , Pediatric Transplant and Cellular Therapy Consortium RESILIENT Conference on Pediatric Chronic Graft-Versus-Host Disease Survivorship After Hematopoietic Cell Transplantation: Part I. Phases of Chronic GVHD, Supportive Care, and Systemic Therapy Discontinuation. *Transplant Cell Ther*, 2025. 31(2): p. 69 e1–69 e18.
10. Hillen U, et al. , Consensus on performing skin biopsies, laboratory workup, evaluation of tissue samples and reporting of the results in patients with suspected cutaneous graft-versus-host disease. *J Eur Acad Dermatol Venereol*, 2015. 29(5): p. 948–54. [PubMed: 25265987]
11. Saad A, et al. , Hematopoietic Cell Transplantation, Version 2.2020, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*, 2020. 18(5): p. 599–634. [PubMed: 32519831]
12. Alvarado SM, et al. , Histopathologic Analysis of Chronic Cutaneous Graft-Versus-Host Disease. *Am J Dermatopathol*, 2024. 46(11): p. 755–760. [PubMed: 39288748]
13. Shulman HM, et al. , NIH Consensus development project on criteria for clinical trials in chronic graft-versus-host disease: II. The 2014 Pathology Working Group Report. *Biol Blood Marrow Transplant*, 2015. 21(4): p. 589–603. [PubMed: 25639770]
14. Yamada H, et al. , Immunohistology of skin and oral biopsies in graft-versus-host disease after bone marrow transplantation and cytokine therapy. *J Allergy Clin Immunol*, 1997. 100(6 Pt 2): p. S73–6. [PubMed: 9440549]
15. Kim J, et al. , Clinical characteristics and histopathologic changes of morphea: A single-center, retrospective study of 137 patients. *J Am Acad Dermatol*, 2021. 85(1): p. 105–113. [PubMed: 33188871]
16. Showalter K and Gordon JK, Skin Histology in Systemic Sclerosis: a Relevant Clinical Biomarker. *Curr Rheumatol Rep*, 2020. 23(1): p. 3. [PubMed: 33244633]
17. Hogenes MCH, et al. , Clinical versus histological grading in the assessment of cutaneous graft versus host disease. *Eur J Med Res*, 2019. 24(1): p. 19. [PubMed: 30971293]
18. Wolff D, et al. , National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: IV. The 2020 Highly morbid forms report. *Transplant Cell Ther*, 2021. 27(10): p. 817–835. [PubMed: 34217703]
19. Baumrin E, et al. , Prognostic Value of Cutaneous Disease Severity Estimates on Survival Outcomes in Patients With Chronic Graft-vs-Host Disease. *JAMA Dermatol*, 2023. 159(4): p. 393–402. [PubMed: 36884224]

20. Curtis LM, et al. , NIH response criteria measures are associated with important parameters of disease severity in patients with chronic GVHD. *Bone Marrow Transplant*, 2014. 49(12): p. 1513–20. [PubMed: 25153693]
21. Inamoto Y, et al. , Incidence, risk factors, and outcomes of sclerosis in patients with chronic graft-versus-host disease. *Blood*, 2013. 121(25): p. 5098–103. [PubMed: 23547053]
22. Markova A, et al. , Topical Ruxolitinib for Chronic Cutaneous GvHD: Promising Results of a Phase 2 Clinical Trial. *Blood*, 2023. 142(Supplement 1): p. 777–777. [PubMed: 36315909]
23. Markova A, et al. , Phase 2 Trial of Ruxolitinib Cream for Chronic Cutaneous GvHD. *Transplantation and Cellular Therapy*, Official Publication of the American Society for Transplantation and Cellular Therapy, 2024. 30(2): p. S1.
24. Belum VR, et al. , A prospective, randomized, double-blinded, split-face/chest study of prophylactic topical dapsone 5% gel versus moisturizer for the prevention of cetuximab-induced acneiform rash. *J Am Acad Dermatol*, 2017. 77(3): p. 577–579. [PubMed: 28807114]
25. Choi CJ and Nghiem P, Tacrolimus ointment in the treatment of chronic cutaneous graft-vs-host disease: a case series of 18 patients. *Arch Dermatol*, 2001. 137(9): p. 1202–6. [PubMed: 11559218]
26. Elad S, et al. , Topical tacrolimus--a novel treatment alternative for cutaneous chronic graft-versus-host disease. *Transpl Int*, 2003. 16(9): p. 665–70. [PubMed: 12768230]
27. Cuvelier GDE, et al. , Benefits and challenges with diagnosing chronic and late acute GVHD in children using the NIH consensus criteria. *Blood*, 2019. 134(3): p. 304–316. [PubMed: 31043425]
28. Rumpf RW, et al. , Comparison of the Lund and Browder table to computed tomography scan three-dimensional surface area measurement for a pediatric cohort. *J Surg Res*, 2018. 221: p. 275–284. [PubMed: 29229139]
29. Li SC and Zheng RJ, Overview of Juvenile localized scleroderma and its management. *World J Pediatr*, 2020. 16(1): p. 5–18. [PubMed: 31786801]
30. Pham C, Collier Z, and Gillenwater J, Changing the Way We Think About Burn Size Estimation. *J Burn Care Res*, 2019. 40(1): p. 1–11. [PubMed: 30247559]
31. DuBois D DE, Fifth paper the measurement of the surface area of man. *Archives of Internal Medicine*, 1915. 15(5_2): p. 868–81.
32. Choi J, et al. , Practical Computer Vision Application to Compute Total Body Surface Area Burn: Reappraising a Fundamental Burn Injury Formula in the Modern Era. *JAMA Surg*, 2022. 157(2): p. 129–135. [PubMed: 34817552]
33. Wurzer P, et al. , Smartphone applications in burns. *Burns*, 2015. 41(5): p. 977–89. [PubMed: 25655039]
34. Chong HP, et al. , A comparison study of methods for estimation of a burn surface area: Lund and Browder, e-burn and Mersey Burns. *Burns*, 2020. 46(2): p. 483–489. [PubMed: 31561927]
35. Colson CD, et al. , EasyTBSA as a method for calculating total body surface area burned: a validation study. *Emerg Med J*, 2023. 40(4): p. 279–284. [PubMed: 36639224]
36. Huang S, et al. , A systematic review of machine learning and automation in burn wound evaluation: A promising but developing frontier. *Burns*, 2021. 47(8): p. 1691–1704. [PubMed: 34419331]
37. McNeil A, et al. , Artificial intelligence recognition of cutaneous chronic graft-versus-host disease by a deep learning neural network. *Br J Haematol*, 2022. 197(6): p. e69–e72. [PubMed: 35322873]
38. Ray WC, et al. , A 1% TBSA Chart Reduces Math Errors While Retaining Acceptable First-Estimate Accuracy. *J Burn Care Res*, 2022. 43(3): p. 665–678. [PubMed: 34665849]
39. Wachtel TL, et al. , The inter-rater reliability of estimating the size of burns from various burn area chart drawings. *Burns*, 2000. 26(2): p. 156–70. [PubMed: 10716359]
40. Murari A and Singh KN, Lund and Browder chart-modified versus original: a comparative study. *Acute Crit Care*, 2019. 34(4): p. 276–281. [PubMed: 31795625]
41. Kamolz LP, et al. , Smartphones and burn size estimation: “Rapid Burn Assessor”. *Ann Burns Fire Disasters*, 2014. 27(2): p. 101–4. [PubMed: 26170784]
42. Knaysi GA, Crikelair GF, and Cosman B, The role of nines: its history and accuracy. *Plast Reconstr Surg*, 1968. 41(6): p. 560–3. [PubMed: 5654897]

43. Wallace AB, The exposure treatment of burns. *Lancet*, 1951. 1(6653): p. 501–4. [PubMed: 14805109]
44. Ramsay B and Lawrence CM, Measurement of involved surface area in patients with psoriasis. *Br J Dermatol*, 1991. 124(6): p. 565–70. [PubMed: 2064940]
45. Rhodes J, Clay C, and Phillips M, The surface area of the hand and the palm for estimating percentage of total body surface area: results of a meta-analysis. *Br J Dermatol*, 2013. 169(1): p. 76–84. [PubMed: 23448271]
46. Perry RJ, et al. , Determining the approximate area of a burn: an inconsistency investigated and re-evaluated. *BMJ*, 1996. 312(7042): p. 1338. [PubMed: 8646048]
47. Agarwal P and Sahu S, Determination of hand and palm area as a ratio of body surface area in Indian population. *Indian J Plast Surg*, 2010. 43(1): p. 49–53. [PubMed: 20924450]
48. Berry MG, Evison D, and Roberts AH, The influence of body mass index on burn surface area estimated from the area of the hand. *Burns*, 2001. 27(6): p. 591–4. [PubMed: 11525853]
49. Martin PJ, et al. , Organ Changes Associated with Provider-Assessed Responses in Patients with Chronic Graft-versus-Host Disease. *Biol Blood Marrow Transplant*, 2019. 25(9): p. 1869–1874. [PubMed: 31085305]
50. Baumrin E, et al. , Patient-Reported Outcomes and Mortality in Cutaneous Chronic Graft-vs-Host Disease. *JAMA Dermatol*, 2024. 160(4): p. 393–401. [PubMed: 38416506]
51. Jacobsohn DA, et al. , Correlation between NIH composite skin score, patient-reported skin score, and outcome: results from the Chronic GVHD Consortium. *Blood*, 2012. 120(13): p. 2545–52; quiz 2774. [PubMed: 22773386]
52. Martires KJ, et al. , Sclerotic-type chronic GVHD of the skin: clinical risk factors, laboratory markers, and burden of disease. *Blood*, 2011. 118(15): p. 4250–7. [PubMed: 21791415]
53. Hamilton BK, et al. , Disability Related to Chronic Graft-versus-Host Disease. *Biol Blood Marrow Transplant*, 2020. 26(4): p. 772–777. [PubMed: 31669175]
54. Lee SJ, et al. , Correlation of Patient-Reported Outcomes with Clinical Organ Responses: Data from the Belumosudil Chronic Graft-versus-Host Disease Studies. *Transplant Cell Ther*, 2022. 28(10): p. 700 e1–700 e6.
55. Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims 2009 December 2009; Available from: <https://www.fda.gov/media/77832/download>.
56. de Vere Hunt I, et al. , “Is this the GVHD?” A qualitative exploration of quality of life issues in individuals with graft-versus-host disease following allogeneic stem cell transplant and their experiences of a specialist multidisciplinary bone marrow transplant service. *Health Qual Life Outcomes*, 2021. 19(1): p. 11. [PubMed: 33413414]
57. Parisek M, et al. , “This Graft-vs.-Host Disease Determines My Life. That’s It.”-A Qualitative Analysis of the Experiences and Needs of Allogeneic Hematopoietic Stem Cells Transplantation Survivors in Germany. *Front Public Health*, 2021. 9: p. 687675. [PubMed: 34277549]
58. Pidala JA, et al. , Comparison of treatment response measures in cutaneous sclerosis after allogeneic hematopoietic cell transplantation. *Blood Adv*, 2024. 8(17): p. 4651–4657. [PubMed: 38885484]
59. Inamoto Y, et al. , Assessment of joint and fascia manifestations in chronic graft-versus-host disease. *Arthritis Rheumatol*, 2014. 66(4): p. 1044–52. [PubMed: 24757155]
60. Merkel EC, Mitchell SA, and 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): p. 752–758. [PubMed: 26751003]
61. Jensen RE, et al. , United States Population-Based Estimates of Patient-Reported Outcomes Measurement Information System Symptom and Functional Status Reference Values for Individuals With Cancer. *J Clin Oncol*, 2017. 35(17): p. 1913–1920. [PubMed: 28426375]
62. Lee JH, et al. , Randomized trial of myeloablative conditioning regimens: busulfan plus cyclophosphamide versus busulfan plus fludarabine. *J Clin Oncol*, 2013. 31(6): p. 701–9. [PubMed: 23129746]
63. Ware JE Jr., SF-36 health survey update. *Spine (Phila Pa 1976)*, 2000. 25(24): p. 3130–9. [PubMed: 11124729]

64. Kilgour JM, et al. , Systematic Review of Patient-Reported Outcome Measures in Graft-versus-Host Disease. *Biol Blood Marrow Transplant*, 2020. 26(5): p. e113–e127. [PubMed: 32028026]
65. Finlay AY and Khan GK, Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. *Clin Exp Dermatol*, 1994. 19(3): p. 210–6. [PubMed: 8033378]
66. Chren MM, et al. , Skindex, a quality-of-life measure for patients with skin disease: reliability, validity, and responsiveness. *J Invest Dermatol*, 1996. 107(5): p. 707–13. [PubMed: 8875954]
67. Ogunsanya ME, et al. , Validation and reliability of a disease-specific quality-of-life measure in patients with cutaneous lupus erythematosus. *Br J Dermatol*, 2019. 180(6): p. 1430–1437. [PubMed: 30637718]
68. Zigler CK, et al. , Cross-sectional quantitative validation of the paediatric Localized Scleroderma Quality of Life Instrument (LoSQI): A disease-specific patient-reported outcome measure. *J Eur Acad Dermatol Venereol*, 2023. 37(7): p. 1406–1414. [PubMed: 36950970]
69. Arai S, et al. , A Randomized Phase II Crossover Study of Imatinib or Rituximab for Cutaneous Sclerosis after Hematopoietic Cell Transplantation. *Clin Cancer Res*, 2016. 22(2): p. 319–27. [PubMed: 26378033]
70. Baumrin E, et al. , Development of the Lee Symptom Scale-Skin Sclerosis for chronic GVHD-associated sclerosis. *Blood*, 2025. 145(12): p. 1321–1332. [PubMed: 39808799]
71. Baumrin E, et al. , Outcomes of Importance to Patients Living With Cutaneous Chronic Graft-vs-Host Disease. *JAMA Dermatol*, 2025. 161(3): p. 281–290. [PubMed: 39745736]
72. Aletaha D, et al. , Definition of treatment response in rheumatoid arthritis based on the simplified and the clinical disease activity index. *Ann Rheum Dis*, 2012. 71(7): p. 1190–6. [PubMed: 22454398]
73. Khanna D, et al. , The American College of Rheumatology Provisional Composite Response Index for Clinical Trials in Early Diffuse Cutaneous Systemic Sclerosis. *Arthritis Rheumatol*, 2016. 68(2): p. 299–311. [PubMed: 26808827]
74. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labelling Claims. 2009 [cited 2024 January 11]; Available from: <https://www.fda.gov/media/77832/download>.
75. Lafyatis R, et al. , Inhibition of beta-Catenin Signaling in the Skin Rescues Cutaneous Adipogenesis in Systemic Sclerosis: A Randomized, Double-Blind, Placebo-Controlled Trial of C-82. *J Invest Dermatol*, 2017. 137(12): p. 2473–2483. [PubMed: 28807667]
76. Cardones AR, et al. , Quantifying Skin Stiffness in Graft-Versus-Host Disease, Morphea, and Systemic Sclerosis Using Acoustic Radiation Force Impulse Imaging and Shear Wave Elastography. *J Invest Dermatol*, 2021. 141(4): p. 924–927 e2. [PubMed: 32950506]
77. Horger M, et al. , Musculocutaneous chronic graft-versus-host disease: MRI follow-up of patients undergoing immunosuppressive therapy. *AJR Am J Roentgenol*, 2009. 192(5): p. 1401–6. [PubMed: 19380569]
78. Neid T, et al. , [Sclerodermiform chronic graft-versus-host disease after allogeneic peripheral blood stem-cell transplantation]. *Dtsch Med Wochenschr*, 2009. 134(21): p. 1106–9. [PubMed: 19437374]
79. Osmola-Mankowska A, et al. , “Assessment of chronic sclerodermoid Graft-versus-Host Disease patients, using 20 MHz high-frequency ultrasonography and cutometer methods”. *Skin Res Technol*, 2013. 19(1): p. e417–22. [PubMed: 22882597]
80. Sauter AW, et al. , Imaging findings and therapy response monitoring in chronic sclerodermatous graft-versus-host disease: preliminary data of a simultaneous PET/MRI approach. *Clin Nucl Med*, 2013. 38(8): p. e309–17. [PubMed: 23455525]
81. Ghosh S, et al. , Interrater reproducibility of the Myoton and durometer devices to quantify sclerotic chronic graft-versus-host disease. *Arch Dermatol Res*, 2023. 315(9): p. 2545–2554. [PubMed: 37227518]
82. Ozturk O, et al. , Performance of myotonometer in the assessment of skin involvement in systemic sclerosis. *Clin Rheumatol*, 2024. 43(2): p. 695–705. [PubMed: 38114818]
83. Farhadfar N, et al. , Reproducibility and repeatability of the Myoton to quantify sclerotic chronic graft-versus-host disease. *Blood Adv*, 2025.

84. Chen F, et al. , Non-invasive measurement of sclerosis in cutaneous cGVHD patients with the handheld device Myoton: a cross-sectional study. *Bone Marrow Transplant*, 2019. 54(4): p. 616–619. [PubMed: 30287938]
85. Baker LX, et al. , Optimal biomechanical parameters for measuring sclerotic chronic graft-versus-host disease. *JID Innov*, 2021. 1(3).
86. Baker LX, et al. , Longitudinal tracking of skin dynamic stiffness to quantify evolution of sclerosis in chronic graft-versus-host disease. *Bone Marrow Transplant*, 2021. 56(4): p. 989–991. [PubMed: 33279941]
87. Clark J, et al. , Magnetic resonance imaging in sclerotic-type chronic graft-vs-host disease. *Arch Dermatol*, 2009. 145(8): p. 918–22. [PubMed: 19687424]
88. Baird K, 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): p. 1083–90. [PubMed: 25771402]
89. Keurentjes AJ, Jakasa I, and Kezic S, Research Techniques Made Simple: Stratum Corneum Tape Stripping. *J Invest Dermatol*, 2021. 141(5): p. 1129–1133 e1. [PubMed: 33888213]
90. Kim BE, et al. , Side-by-Side Comparison of Skin Biopsies and Skin Tape Stripping Highlights Abnormal Stratum Corneum in Atopic Dermatitis. *J Invest Dermatol*, 2019. 139(11): p. 2387–2389 e1. [PubMed: 31176708]
91. Tsoi LC, et al. , Noninvasive Tape-Stripping with High-Resolution RNA Profiling Effectively Captures a Preinflammatory State in Nonlesional Psoriatic Skin. *J Invest Dermatol*, 2022. 142(6): p. 1587–1596 e2. [PubMed: 34808239]
92. Balbir-Gurman A, et al. , Non-invasive measurement of biomechanical skin properties in systemic sclerosis. *Ann Rheum Dis*, 2002. 61(3): p. 237–41. [PubMed: 11830429]
93. Pierard GE, Hermanns-Le T, and Pierard-Franchimont C, Scleroderma: skin stiffness assessment using the stress-strain relationship under progressive suction. *Expert Opin Med Diagn*, 2013. 7(2): p. 119–25. [PubMed: 23530882]
94. Pierard GE, et al. , In vivo evaluation of the skin tensile strength by the suction method: pilot study coping with hysteresis and creep extension. *ISRN Dermatol*, 2013. 2013: p. 841217. [PubMed: 23986871]
95. Hiebert P, et al. , A lysyl oxidase-responsive collagen peptide illuminates collagen remodeling in wound healing. *Matrix Biol*, 2024. 128: p. 11–20. [PubMed: 38382767]
96. Arnoldini S, et al. , Novel peptide probes to assess the tensional state of fibronectin fibers in cancer. *Nat Commun*, 2017. 8(1): p. 1793. [PubMed: 29176724]
97. Pinal-Fernandez I, et al. , Pathological autoantibody internalisation in myositis. *Ann Rheum Dis*, 2024.
98. Pinal-Fernandez I, et al. , Transcriptomic profiling reveals distinct subsets of immune checkpoint inhibitor induced myositis. *Ann Rheum Dis*, 2023. 82(6): p. 829–836. [PubMed: 36801811]
99. Liu X, et al. , Baseline Photos and Confident Annotation Improve Automated Detection of Cutaneous Graft-Versus-Host Disease. *Clin Hematol Int*, 2021. 3(3): p. 108–115. [PubMed: 34820616]
100. Tkaczyk ER, et al. , Overcoming human disagreement assessing erythematous lesion severity on 3D photos of chronic graft-versus-host disease. *Bone Marrow Transplant*, 2018. 53(10): p. 1356–1358. [PubMed: 29740182]
101. Parks K, et al. , Non-Expert Markings of Active Chronic Graft-Versus-Host Disease Photographs: Optimal Metrics of Training Effects. *J Digit Imaging*, 2023. 36(1): p. 373–378. [PubMed: 36344635]
102. McNeil AJ, et al. , Crowdsourcing Skin Demarcations of Chronic Graft-Versus-Host Disease in Patient Photographs: Training Versus Performance Study. *JMIR Dermatol*, 2023. 6: p. e48589. [PubMed: 38147369]
103. Bruggen MC, et al. , Diverse T-cell responses characterize the different manifestations of cutaneous graft-versus-host disease. *Blood*, 2014. 123(2): p. 290–9. [PubMed: 24255916]
104. Strong Rodrigues K, et al. , Cutaneous Graft-Versus-Host Disease: Diagnosis and Treatment. *Am J Clin Dermatol*, 2018. 19(1): p. 33–50. [PubMed: 28656563]

105. Strobl J, et al. , Diverse macrophage populations contribute to distinct manifestations of human cutaneous graft-versus-host disease. *Br J Dermatol*, 2024. 190(3): p. 402–414. [PubMed: 38010706]
106. Lehman JS, et al. , Differential expression of interferon-induced genes and other tissue-based biomarkers in acute graft-versus-host disease vs. lupus erythematosus in skin. *Clin Exp Dermatol*, 2019. 44(4): p. e81–e88. [PubMed: 30280423]
107. Wenzel J, et al. , CXCR3 <-> ligand-mediated skin inflammation in cutaneous lichenoid graft-versus-host disease. *J Am Acad Dermatol*, 2008. 58(3): p. 437–42. [PubMed: 18280341]
108. Wenzel J and Tuting T, An IFN-associated cytotoxic cellular immune response against viral, self-, or tumor antigens is a common pathogenetic feature in “interface dermatitis”. *J Invest Dermatol*, 2008. 128(10): p. 2392–402. [PubMed: 18418411]
109. Aquino TM, Calvarido MG, and North JP, Interleukin 36 expression in psoriasis variants and other dermatologic diseases with psoriasis-like histopathologic features. *J Cutan Pathol*, 2022. 49(2): p. 123–132. [PubMed: 34346097]
110. DaCosta A, et al. , Differentiating psoriasis from eczema with interleukin-36 gamma immunohistochemistry on skin biopsy specimens: A retrospective, single-center validation and cohort study. *J Am Acad Dermatol*, 2025.
111. Agarwal A, et al. , Off-Label Use of Janus Kinase Inhibitors in Inflammatory Cutaneous Diseases. *J Drugs Dermatol*, 2023. 22(12): p. 1183–1190. [PubMed: 38051858]
112. Wu JM, et al. , CD20, AIF-1, and TGF-beta in graft-versus-host disease: a study of mRNA expression in histologically matched skin biopsies. *Mod Pathol*, 2010. 23(5): p. 720–8. [PubMed: 20190732]
113. Paczesny S, 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): p. 780–92. [PubMed: 25644957]
114. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource 2016; Available from: <https://www.ncbi.nlm.nih.gov/books/NBK326791/>.
115. Nash RA, et al. , High-dose immunosuppressive therapy and autologous hematopoietic cell transplantation for severe systemic sclerosis: long-term follow-up of the US multicenter pilot study. *Blood*, 2007. 110(4): p. 1388–96. [PubMed: 17452515]
116. Fleming JN, et al. , Cutaneous chronic graft-versus-host disease does not have the abnormal endothelial phenotype or vascular rarefaction characteristic of systemic sclerosis. *PLoS One*, 2009. 4(7): p. e6203. [PubMed: 19587802]
117. Sharma R, et al. , Belumosudil reduces oral chronic graft-versus-host disease tissue inflammation and fibrosis: a ROCKstar companion study. *Blood Adv*, 2025.
118. Garuffo L, et al. , The Applications of Machine Learning in the Management of Patients Undergoing Stem Cell Transplantation: Are We Ready? *Cancers (Basel)*, 2025. 17(3).
119. Kitko CL, et al. , National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: IIa. The 2020 Clinical Implementation and Early Diagnosis Working Group Report. *Transplant Cell Ther*, 2021. 27(7): p. 545–557. [PubMed: 33839317]
120. Williams KM, et al. , National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2020 Etiology and Prevention Working Group Report. *Transplant Cell Ther*, 2021. 27(6): p. 452–466. [PubMed: 33877965]
121. Inamoto Y, et al. , Dickkopf-related protein 3 is a novel biomarker for chronic GVHD after allogeneic hematopoietic cell transplantation. *Blood Adv*, 2020. 4(11): p. 2409–2417. [PubMed: 32492155]
122. Logan BR, et al. , Validated graft-specific biomarkers identify patients at risk for chronic graft-versus-host disease and death. *J Clin Invest*, 2023. 133(15).
123. Krupnik VE, et al. , Functional and structural diversity of the human Dickkopf gene family. *Gene*, 1999. 238(2): p. 301–13. [PubMed: 10570958]
124. Whittle R and Taylor PC, Circulating B-cell activating factor level predicts clinical response of chronic graft-versus-host disease to extracorporeal photopheresis. *Blood*, 2011. 118(24): p. 6446–9. [PubMed: 22021372]

125. Sarantopoulos S, et al. , High levels of B-cell activating factor in patients with active chronic graft-versus-host disease. *Clin Cancer Res*, 2007. 13(20): p. 6107–14. [PubMed: 17947475]
126. Schulz E, et al. , Red cell distribution width as a new prognostic biomarker in refractory chronic graft-versus-host disease. *Haematologica*, 2024. 109(1): p. 298–302. [PubMed: 37584292]
127. Hakim FT, et al. , Upregulation of IFN-Inducible and Damage-Response Pathways in Chronic Graft-versus-Host Disease. *J Immunol*, 2016. 197(9): p. 3490–3503. [PubMed: 27694491]
128. Zorn E, et al. , Reduced frequency of FOXP3+ CD4+CD25+ regulatory T cells in patients with chronic graft-versus-host disease. *Blood*, 2005. 106(8): p. 2903–11. [PubMed: 15972448]
129. Imanguli MM, et al. , Comparative analysis of FoxP3(+) regulatory T cells in the target tissues and blood in chronic graft versus host disease. *Leukemia*, 2014. 28(10): p. 2016–27. [PubMed: 24577531]
130. Boiko JR, et al. , Defining pathogenic IL-17 and CSF-1 gene expression signatures in chronic graft-versus-host disease. *Blood*, 2025. 145(19): p. 2214–2228. [PubMed: 39977705]
131. Bayle J, et al. , Increased expression of Wnt2 and SFRP4 in Tsk mouse skin: role of Wnt signaling in altered dermal fibrillin deposition and systemic sclerosis. *J Invest Dermatol*, 2008. 128(4): p. 871–81. [PubMed: 17943183]
132. Tinazzi I, et al. , SFRP4 Expression Is Linked to Immune-Driven Fibrotic Conditions, Correlates with Skin and Lung Fibrosis in SSc and a Potential EMT Biomarker. *J Clin Med*, 2021. 10(24).
133. Rosenstein RK, et al. , Identification of Fibroinflammatory and Fibrotic Transcriptomic Subsets of Human Cutaneous Sclerotic Chronic Graft-Versus-Host Disease. *JID Innov*, 2024. 4(2): p. 100246. [PubMed: 38357212]
134. Li X, et al. , Overexpression of SerpinE2/protease nexin-1 Contribute to Pathological Cardiac Fibrosis via increasing Collagen Deposition. *Sci Rep*, 2016. 6: p. 37635. [PubMed: 27876880]
135. Bruggen MC, et al. , Epidermal elafin expression is an indicator of poor prognosis in cutaneous graft-versus-host disease. *J Invest Dermatol*, 2015. 135(4): p. 999–1006. [PubMed: 25405322]
136. Kim M, et al. , Sclerotic-Type Cutaneous Chronic Graft-Versus-Host Disease Exhibits Activation of T Helper 1 and OX40 Cytokines. *J Invest Dermatol*, 2024. 144(3): p. 563–572 e9. [PubMed: 37742913]
137. Zouali H, et al. , RNA sequencing of chronic GVHD skin lesions defines shared and unique inflammatory pathways characterizing lichen planus and morphea. *Blood Adv*, 2022. 6(9): p. 2805–2811. [PubMed: 35008096]
138. Croudace JE, et al. , Chemokine-mediated tissue recruitment of CXCR3+ CD4+ T cells plays a major role in the pathogenesis of chronic GVHD. *Blood*, 2012. 120(20): p. 4246–55. [PubMed: 23012327]
139. Lopes SM, et al. , Should we use extracorporeal photopheresis more often? Evidence from graft-versus-host disease patients monitored with Treg as a biomarker. *Future Sci OA*, 2020. 6(8): p. FSO623. [PubMed: 32983569]
140. Murari A, A modified Lund and Browder chart. *Indian J Plast Surg*, 2017. 50(2): p. 220–221. [PubMed: 29343901]

Case Study: Illustrating the need to improve the skin-specific response criteria for patients with epidermal cutaneous chronic GVHD.

A 68-year-old man presented 10 months post-transplant with severe cGVHD, refractory to systemic steroids. At baseline, he had an epidermal cutaneous cGVHD with lichen planus-like features (90% BSA). His extensive BSA using the 2014 NIH criteria was scored as a 3. Accordingly, his overall global severity score for cGVHD was 3. He initiated an approved systemic therapy for cGVHD. At 3 months, the BSA involvement was significantly reduced to 55% of active BSA with 35% of hyperpigmented BSA resulting in Stable Disease by the NIH 2014 Criteria, but Partial Response in Epidermal cGVHD by the proposed 2025 response assessment (Figure 3). By 6 months, the patient’s BSA further improved to 20%, and Partial Response was detected on both response assessments.

Using the proposed 2025 changes, this patient’s skin response would be captured as a PR sooner due to a ≥20% BSA improvement and a minimum of absolute 10% reduction from baseline BSA. This exemplifies the potential benefit of the proposed revision and expanded criteria for skin response, which would better capture meaningful improvements in patients with epidermal skin GVHD involvement but will require further validation in the clinical trial setting.

Study Timepoint		Skin Score (NIH 2014 criteria)	BSA (%) Epidermal changes	BSA (%) Sclerosis
Baseline	✓ Lichen planus-like features	3	90	0
3 months	✓ Lichen planus-like features	3	55	0
6 months	✓ Lichen planus-like features	2	20	0

Responses at 6- months per various parameters contributing to meaningful skin-response and formal 2025 skin response scoring. Green=Response; Yellow=Stable disease.

		Current Skin Response 2014 Criteria	Proposed Epidermal Skin Response 2025 Criteria	Proposed Sclerosis Skin Response 2025 Criteria
3 months		Stable Disease (SD)	Partial Response (PR)	N/A
6 months		Partial Response (PR)	Partial Response (PR)	N/A

Final skin-specific response per current, formal 2014 criteria versus proposed 2025 criteria and the contributing factors to each response. Arrows highlight contributing factors to each response criteria. BSA, Body surface area.

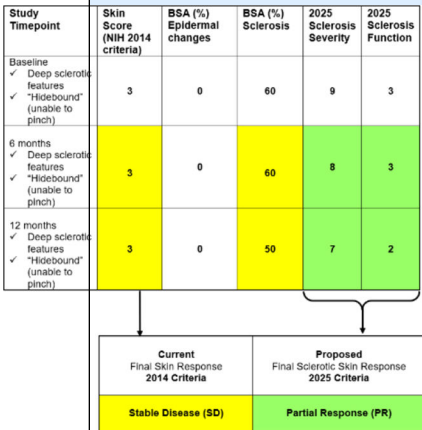
Figure 3.
Case Study Illustration of:

Case Study: Illustrating the need to improve the skin-specific response criteria for patients with sclerotic cutaneous chronic GVHD.

A 50-year-old woman presented 5 years post-transplant with severe cGVHD of the skin, mouth, and joint/fascia, refractory to 11 prior lines of systemic therapy. At baseline, she had a mix of superficial and deep sclerotic features (60% BSA) with areas of hidebound, rippled skin, hyperpigmented morphea-like changes, a superficial ulceration on the left foot, and pruritus. The presence of deep-seated sclerosis classified her skin involvement using the 2014 NIH criteria as a score 3. After 6 months of treatment on a systemic clinical trial, the BSA involvement was unchanged, but she reported significant skin softening and increased ROM. In addition, she had resolution of pruritus and healing of the leg ulcer. At 12 months, BSA involvement of sclerosis decreased from 60% to 50%. At the end of treatment visit, the clinician noted continued improvement in skin softening in the patient’s arms, hips, and legs with less hyperpigmentation, increased body hair growth, and improved ability to brush her hair, bend over to tie her shoes, and reach for items on higher shelves. The exploratory provider-reported skin-tightening severity scale (0–10 scale) also showed a 2-point reduction from baseline with treatment.

Despite improvement in %BSA sclerosis, and provider-reported skin-specific severity (Figure 4), the formal scoring of the skin remained a ‘3’ due to continued presence of hidebound/deep seated sclerosis, and her skin organ response remained ‘Stable Disease’. The skin score did not capture other meaningful improvement such as resolution of diffuse pruritus, skin ulceration, and cramping – which all further impacted the patient’s sclerosis severity and function.

Using the proposed 2025 changes, this patient’s skin response would be captured as a PR due to her fulfilling the criteria of ‘no worsening in severity or function of sclerosis’ AND the two following criteria (a) Sclerosis Severity improvement of 2 points and (b) Sclerosis Function (as appreciated by patient’s performance status and functional assessments) by one point. This exemplifies the potential benefit of the proposed revision and expanded criteria for skin response, which would better capture meaningful improvements for our sclerotic skin GVHD patients but will require further validation in the clinical trial setting.



Responses at 6- and 12-months per various parameters contributing to meaningful skin-response and formal 2025 skin response scoring. Green: Response; Yellow: Stable disease.

Final skin-specific response per current, formal 2014 criteria versus proposed 2025 criteria and the contributing factors to each response. Arrows highlight contributing factors to each response criteria. BSA, Body surface area.

Figure 4.

Case Study Illustration of:

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- NIH Task Force refines response criteria for chronic cutaneous graft-versus-host disease
- Body surface area scoring separated for epidermal vs sclerotic skin involvement
- Two new clinician scales measure sclerosis severity and qualitative features
- Revised sclerosis assessments improve sensitivity to clinically meaningful change

	SCORE 0	SCORE 1	SCORE 2	SCORE 3
SKIN <u>GVHD features to be scored by BSA:</u> Check all that apply: ☐ Maculopapular rash / erythema ☐ Lichen planus-like features ☐ Sclerotic features ☐ Papulosquamous lesions or ichthyosis ☐ Keratosis pilaris-like	☐ No BSA involved	☐ 1-18% BSA	☐ 19-50% BSA	☐ >50% BSA
☐ Abnormality present but explained entirely by non-GVHD documented cause (specify): _____				
SKIN FEATURES SCORE:	☐ No sclerotic features		☐ Superficial sclerotic features "not hidebound" (able to pinch)	Check all that apply: ☐ Deep sclerotic features ☐ "Hidebound" (unable to pinch) ☐ Impaired mobility ☐ Ulceration
If skin features score = 3, BSA% of non-moveable sclerosis/fasciitis _____ How would you rate the severity of this patient's skin and/or joint tightening on the following scale, where 0 is not at all severe and 10 is the most severe symptoms possible: 0 1 2 3 4 5 6 7 8 9 10 Symptoms not at all severe Most severe symptoms possible				

Figure 1A.
 2014 NIH CHRONIC GVHD ACTIVITY ASSESSMENT- CLINICIAN (FORM A [7])

Organ	Complete Response	Partial Response	Progression
Skin	NIH Skin Score 0 after previous involvement	Decrease in NIH Skin Score by 1 or more points	Increase in NIH Skin Score by 1 or more points, except 0 to 1

Figure 1B.
2014 NIH Response determination for chronic GVHD clinical trials based on clinician assessments[7]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

2025 Chronic Cutaneous GVHD Response Assessment: Age > 15 Years

GVHD features to be assessed by BSA:
(Check all that apply)

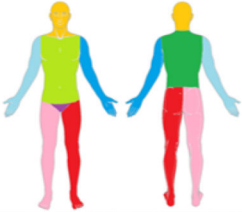
Epidermal Skin

☐ Maculopapular rash
☐ Lichen planus-like features
☐ Papulosquamous lesions or ichthyosis
☐ Keratosis pilaris-like

Sclerotic Skin

☐ Superficial sclerotic features "not hidebound" (able to pinch; e.g., lichen sclerosus-like, morphea-like)
☐ Deep sclerotic features
☐ "Hidebound" (unable to pinch)
☐ Sclerosis-associated edema +/- erythema
☐ GVHD-related ulceration

BSA calculator



Age >15 years	Epidermal changes		Sclerosis			
	% region involved (0-100%)	Multiplier (%)	Total BSA (%)	% region involved (0-100%)	Multiplier (%)	Total BSA (%)
Head/neck/scalp		0.09			0.09	
Anterior torso		0.13			0.13	
Posterior torso		0.13			0.13	
L. upper extrem.		0.10			0.10	
R. upper extrem.		0.10			0.10	
L. lower extrem. (incl. L. buttock)		0.23			0.23	
R. lower extrem. (incl. R. buttock)		0.23			0.23	
Genitalia		0.01			0.01	

If patient has sclerotic features documented in BSA, complete both below:

SCLEROSIS SEVERITY

How would you rate the quality and physical signs of the patient's skin sclerosis on the following scale?

0 No sclerosis
1
2
3
4
5 Moderate sclerosis
6
7
8
9
10 Most severe sclerosis

Qualitative changes include, but are not limited to: Softening of fibrosis, healing of ulcers, improving lymphedema, hair regrowth, skin tightness, abdominal fullness, dimpling/rippling etc.

SCLEROSIS FUNCTION

What is the impact of the cutaneous sclerotic GVHD on the patient's daily function?

0 None
1 Minimal
2 Mild
3 Moderate
4 Severe

Figure 2A.
Proposed 2025 NIH Consensus Skin Task Force Cutaneous Response Assessment

Transplant Cell Ther. Author manuscript; available in PMC 2026 February 13.

Cutaneous cGVHD Response Epidermal and Sclerotic to be assessed separately	Epidermal GVHD Response	Sclerotic GVHD Response
Complete Response	Resolution of active epidermal BSA to 0%	Resolution of sclerotic GVHD to severity scale 0, function scale 0, AND BSA 0%
Partial Response	≥ 20% improvement of epidermal BSA (i.e., ≥0.2 x baseline BSA) AND a minimum of absolute 10% reduction from baseline BSA; and BSA >0%	No worsening in severity or function of sclerosis* AND any 2 of the following: • ≥ 20% improvement of baseline BSA and a minimum absolute 10% improvement from baseline BSA (or resolution to 0%) • ≥ 2 points improvement on clinician severity assessment (0-10) scale • ≥ 1 point improvement on function (0-4) scale
Stable Disease	Not meeting criteria for Response or Progression	Not meeting criteria for Response or Progression
Progression	>20% worsening of epidermal BSA AND a minimum of absolute 10% worsening from baseline	Any 2 of the following: • >20% worsening of baseline sclerosis BSA and a minimum absolute 10% increase from baseline BSA • ≥ 2 points worsening on sclerosis clinician severity assessment (0-10) scale • ≥ 1 point worsening on sclerosis function (0-4) scale

*No worsening in severity by 2 or more points on 0-10 scale or by 1 or more points on function (0-4) scale

Figure 2B.
Proposed 2025 Cutaneous Chronic GVHD Response Determination

Table 1.
Limitations of current cutaneous chronic GVHD response criteria

Problem	Suggestions for Improvement
• %BSA is poorly reproducible, and score categories may over- or under detect clinically meaningful change. Basis: Per the NIH 2014 criteria, extent of skin involvement is scored 0–3 based on 4 categories (None, 1–18%, 19–50%, >50% BSA). The threshold for defining change exceeding measurement error varies by phenotype and is estimated to be approximately 20% [8]. Coarseness of %BSA categories may falsely capture insignificant change (e.g., BSA decrease from 55% (score 3) to 45% (score 2), while failing to capture substantial improvements (e.g., BSA decrease from 90% (score 3) to 55% (score 3) response).	• Score change associated with BSA change should correlate with a clinically meaningful benefit or detriment to the patient.
• Combined scoring of epidermal and sclerotic chronic GVHD phenotypes prevents assessment of treatment response for a specific phenotype. Basis: Current NIH 2014 overall skin score is combined for both epidermal and sclerotic features, using the higher of the two scores for the final NIH skin score: %BSA and skin features. This score is currently used as the primary outcome measure for skin organ-specific scoring in clinical trials. Combined overall BSA includes areas of potential overlap in epidermal and sclerotic features. While epidermal chronic GVHD may improve more rapidly than sclerotic chronic GVHD, the response endpoint for both subtypes is the same.	• Epidermal and sclerotic phenotypes have different anatomic distribution, risk factors, biomarkers, and prognosis; therefore, accurate and separate BSA estimation of epidermal and sclerotic features is needed. • % BSA should be captured accurately, and separately for each disease subtype (i.e., epidermal, sclerotic) at diagnosis and at longitudinal response assessment timepoints, enabling detection of discrepant treatment responses between chronic GVHD phenotypes.
• The current skin features score may not capture the clinician-assessed quality and functional impact of sclerosis [22]. Basis: Current NIH 2014 features score for sclerosis measures: 0: No sclerotic features, 2: superficial sclerotic features, and 3: deep sclerotic features. Deep sclerosis, regardless of activity (i.e., active vs inactive, fixed/irreversible disease), currently receives a score of 3, preventing capture of clinician-perceived severity of evolving sclerosis or functional change.	• Quantitative scales to capture clinician-assessed qualitative disease signs/symptoms and functional changes (e.g., skin softening[22], tightening, decreased edema, body hair regrowth, etc.) in sclerotic chronic GVHD are needed to improve the sensitivity of detecting clinically meaningful responses, particularly for organ-specific clinical trials targeting sclerotic disease. • Development of quantitative and reproducible methods is needed to evaluate the extent and depth of sclerosis, with imaging modalities holding potential for future application.

TABLE 2.
Reported NIH skin chronic GVHD response rates in clinical trials

	Ibrutinib[4]	Belumosudil[2]	Ruxolitinib[3]	Axatilimab[1]
FDA Approval	August 2017	July 2021	Sept 2021	August 2024
Study design	Phase 1b/2	Phase 2 - two dosing cohorts: 200mg PO QD 200mg PO BID	Phase 3, open-label: ruxolitinib vs. best available care (control arm)	Phase 2 - three different dosing cohorts: 0.3mg/kg q2w 1.0mg/kgq2w 3.0mg/kg q4w
Prior lines of therapy (LOT) Eligibility; Median	Eligibility: ≤ 3 LOT; Median: 2 LOT	2–5 LOT 3 LOT	≤ 2 LOT N/A	≥ 2 LOT 4 LOT
Overall response rate	67% (CR: 9%)	74% in QD arm (CR: 5%)	76.4% (vs 60.4% in control arm) (CR: 12.1 vs 6.7%)	74% in 0.3mg/kg q2w (CR: 1.3%)
Cutaneous chronic GVHD among all study patients	34 of 42 (80%)	110 of 132 (83%)	234 of 329 (71%)	193 of 241 (80%)
Skin ORR	88% (CR: N/A)	37% (CR: 16%)	41.2% (vs 15.5% in control arm) (CR: N/A)	26% (CR: 9%)
Skin BSA Score *2005 NIH Criteria	88% with baseline skin involvement had improvement (Eligibility required 25% BSA erythematous rash). The study was conducted before the 2014 NIH consensus criteria and so did not report sclerotic skin feature details or clinician-perceived sclerosis severity.	15 / 110 (13%) had improvement in skin BSA score	Not reported	Reported improvement of % BSA non-moveable sclerosis/fasciitis, 43 patients: 44.2% Improved 51.2% Not improved 4.7% Missing
Skin feature score		11 / 110 (10%) had improvement in skin feature score	Baseline: - Score 1: 19/119 (16%) - Score 2,3: 100/119 (84%) Organ response is reported as aggregate. Does not detail responses of patients with epidermal, movable sclerotic, or fixed sclerotic at baseline.	See above, reported only skin feature score 3 Modified Total Lee Symptom scale reported, 45 patients: 73.3% improved 22.2% not improved 4.4% missing
Clinician-perceived sclerosis severity scale (0 to 10)		2 / 110 (1.8%) had improvement in the clinician-perceived sclerosis severity scale	Not reported	53 patients 66% improved 28.3% not improved 5.7% missing