Challenges in clinical practice using the NIH criteria

Daniel Wolff

Dept. of Medicine III, University Hospital Regensburg
## Disclosures

<table>
<thead>
<tr>
<th>Category</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Support/P.I.</td>
<td>Dr. Falk Pharma</td>
</tr>
<tr>
<td>Employee</td>
<td>No relevant conflicts of interest to declare</td>
</tr>
<tr>
<td>Consultant</td>
<td>Dr. Falk Pharma</td>
</tr>
<tr>
<td>Major Stockholder</td>
<td>No relevant conflicts of interest to declare</td>
</tr>
<tr>
<td>Speakers Bureau</td>
<td>No relevant conflicts of interest to declare</td>
</tr>
<tr>
<td>Honoraria</td>
<td>Novartis, Therakos, Fresenius</td>
</tr>
<tr>
<td>Scientific Advisory Board</td>
<td>No relevant conflicts of interest to declare</td>
</tr>
</tbody>
</table>
Application of NIH consensus in clinical routine: Diagnosis

- NIH criteria for diagnosis provide defined criteria for diagnosis of GVHD
- Differentiation between acute and chronic GVHD independent of time after transplantation achieved
- Diagnostic features confirm the presence of cGVHD – histological confirmation is not required
- Distinctive features confirm chronic GVHD if GVHD is shown in histology
- Indication for histological confirmation provided

- Diagnostic criteria were originally developed for clinical trials – their role in clinical practice remain to be evaluated
Application of NIH consensus in clinical routine: Organ and overall severity grading

- NIH grading has been originally developed to improve validity and comparability of clinical trials
- NIH organ grading can be performed by a transplant physician without external help and therefor can be applied in clinical routine
  - Application of NIH criteria improves documentation of the extend and severity of cGVHD
  - Long term follow up requires standardized documentation (changing care provider, retrospective bias, etc.)
- Permits response assessment
- Defines indication for topical and systemic immunosuppression
- Serves as a check list
- Permits retrospective analyses and register studies
Application of NIH consensus in clinical routine: diagnosis and grading

<table>
<thead>
<tr>
<th>Table 2. Use of NIH consensus criteria and barriers to greater use, and interest in materials and training</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>North and South America</strong> (American group), n (%)</td>
</tr>
<tr>
<td>Response rate</td>
</tr>
</tbody>
</table>

- Familiarity with the NIH consensus criteria
  - **Yes**: 284 (94%) in North and South America, 96 (96%) in Europe, Asia, Australia, Africa
  - **No**: 18 (6%) in North and South America, 4 (4%) in Europe, Asia, Australia, Africa

- If familiar with the NIH consensus criteria (n = 284), comfort with making the diagnosis of chronic GVHD according to NIH criteria
  - **Yes, definitely**: 132 (47%) in North and South America, 40 (40%) in Europe, Asia, Australia, Africa
  - **Yes, somewhat**: 149 (52%) in North and South America, 41 (41%) in Europe, Asia, Australia, Africa
  - **No**: 3 (1%) in North and South America, 6 (6%) in Europe, Asia, Australia, Africa
  - **Missing**: 0 in North and South America, 13 (13%) in Europe, Asia, Australia, Africa

- If familiar with the NIH consensus criteria (n = 284), comfort with calculating the mild, moderate or severe global severity score
  - **Yes, definitely**: 109 (38%) in North and South America, 38 (38%) in Europe, Asia, Australia, Africa
  - **Yes, somewhat**: 159 (56%) in North and South America, 42 (42%) in Europe, Asia, Australia, Africa
  - **No**: 14 (5%) in North and South America, 7 (7%) in Europe, Asia, Australia, Africa
  - **Missing**: 2 (1%) in North and South America, 13 (13%) in Europe, Asia, Australia, Africa

- If familiar with the NIH consensus criteria (n = 284), routine use of NIH criteria for diagnosis and severity in clinical practice
  - **Yes**: 195 (69%) in North and South America, 54 (54%) in Europe, Asia, Australia, Africa
  - **No**: 86 (30%) in North and South America, 33 (33%) in Europe, Asia, Australia, Africa
  - **Missing**: 3 (1%) in North and South America, 13 (13%) in Europe, Asia, Australia, Africa

- NIH grading is applied by a minority of the Tx centers in clinical routine within the EBMT and US

Duarte, R. BMT 2013
Application of NIH consensus in clinical routine: Diagnosis and grading

Diagnosis and Staging of Chronic Graft-versus-Host Disease in the Clinical Practice

Hildegard T. Greinix, Christoph Loddenkemper, Steven Z. Pavletic, Ernst Holler, Gerard Socié, Anita Lawitschka, Jörg Halter, Daniel Wolff

Table 2. Results of Survey on Daily Practice of Diagnosis and Staging of Chronic GVHD

<table>
<thead>
<tr>
<th>Question</th>
<th>No Centers in Agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you use the NIH criteria for diagnosis of cGVHD in daily routine?</td>
<td>11 (85)</td>
</tr>
<tr>
<td>Do you distinguish between classic cGVHD and overlap syndrome?</td>
<td>10 (77)</td>
</tr>
<tr>
<td>Do you agree with the definition of classic cGVHD?</td>
<td>13 (100)</td>
</tr>
<tr>
<td>Do you agree with the definition of overlap syndrome?</td>
<td>12 (92)</td>
</tr>
<tr>
<td>At least 1 diagnostic sign* is necessary for diagnosis of cGVHD.</td>
<td>13 (100)</td>
</tr>
<tr>
<td>When a diagnostic sign* is missing, a distinctive sign with confirmation by lab, radiology, or biopsy is required for diagnosis of cGVHD.</td>
<td>12 (92)</td>
</tr>
<tr>
<td>Do you agree with the diagnostic signs* of cGVHD of the skin?</td>
<td>13 (100)</td>
</tr>
<tr>
<td>Do you agree with the distinctive signs* of cGVHD of the skin?</td>
<td>13 (100)</td>
</tr>
<tr>
<td>Do you document the % of superficial sclerosis of the skin?</td>
<td>8 (61.5)</td>
</tr>
<tr>
<td>Do you document the % of deep sclerosis of the skin?</td>
<td>6 (46)</td>
</tr>
<tr>
<td>Do you document the % of hypo/hyperpigmentation of the skin?</td>
<td>3 (23)</td>
</tr>
<tr>
<td>We perform routinely skin biopsies in cGVHD patients.</td>
<td>5 (38)</td>
</tr>
<tr>
<td>Do you agree with the diagnostic signs* of cGVHD of the oral mucosa?</td>
<td>12 (92)</td>
</tr>
<tr>
<td>Do you agree with the distinctive signs* of cGVHD of the oral mucosa?</td>
<td>11 (85)</td>
</tr>
<tr>
<td>We routinely exclude infections of the oral mucosa.</td>
<td>9 (69)</td>
</tr>
<tr>
<td>Do you agree with the distinctive signs* of cGVHD of the eyes?</td>
<td>12 (92)</td>
</tr>
<tr>
<td>We routinely perform Schirmer tests.</td>
<td>4 (31)</td>
</tr>
<tr>
<td>Symptomatic patients are routinely seen by an ophthalmologist.</td>
<td>13 (100)</td>
</tr>
<tr>
<td>Do you agree with the diagnostic signs* of cGVHD of the genitalia?</td>
<td>11 (85)</td>
</tr>
<tr>
<td>Do you agree with the distinctive signs* of cGVHD of the genitalia?</td>
<td>11 (85)</td>
</tr>
<tr>
<td>Symptomatic patients are routinely seen by a gynecologist.</td>
<td>12 (92)</td>
</tr>
</tbody>
</table>
Mild cGVHD (2 organs only involved with mild grade – no lung involvement):
- symptomatic treatment with topical immunosuppression or steroids alone (no impairment of organ function, improved prognosis in malignant diseases) (Filipovich 2005, Jagasia 2014, Wolff 2010)

Moderate cGVHD (either more than 2 organs involved or moderate involvement or mild lung involvement):
- indication for systemic immunosuppression (steroids or steroids + CNI), topical immunosuppression may be added

Severe cGVHD (severe organ involvement or moderate lung involvement):
- consider combination treatment – impaired prognosis
Solved and unsolved issues within the NIH consensus

• How to categorize a patient with isolated symptoms of acute GVHD after chronic GVHD – *solved within the 2014 update* – regarded as *acute*
  
  ▪ Broad definition of Overlap Syndrome (every type of chronic GVHD showing common symptoms of GVHD) resulted in heterogeneous application of the term
    ➢ Current solution: documentation of any acute component as it is by symptoms but no formal classification – is regarded as cGVHD
  
  ▪ Grading of rare manifestations (arthritis, nephrotic syndrome, polyneuropathia, myasthenia gravis, polymyositis) - *not well defined*
Clinical challenges in NIH grading – skin - diagnosis

- Is histological confirmation in clinical routine required if diagnostic symptoms are lacking?
  - 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 U. Hillen et al JEADV in press 2014: If cGVHD is not confirmed at other sites and suspected cutaneous GVHD lacks diagnostic signs biopsy is recommended
    - Reasons: 7% (Jacobsohn 2001) -16% (Paun 2013) of performed biopsies result in change of treatment
    - Difficulties: expertise of the pathologist, clinical information available to the Pathologist (Ziemer et al JEADV 2013)
Clinical challenges in NIH grading – skin - histology

provided by M. Ziemer
# Clinical challenges in NIH grading – skin grading

<table>
<thead>
<tr>
<th>SKIN†</th>
<th>none</th>
<th>mild</th>
<th>moderate</th>
<th>severe</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SCORE % BSA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GVHD features to be scored by BSA:</strong></td>
<td>□ No BSA involved</td>
<td>□ 1-18% BSA</td>
<td>□ 19-50% BSA</td>
<td>□ &gt;50% BSA</td>
</tr>
<tr>
<td><strong>Check all that applies:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>□ Maculopapular rash/erythema</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>□ Lichen planus-like features</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>□ Sclerotic features</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>□ Papulosquamous lesions or ichthyosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>□ Keratosis pilaris-like GVHD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SKIN FEATURES SCORE:</strong></td>
<td>□ No sclerotic features</td>
<td></td>
<td>□ Superficial sclerotic features “not hidebound” (able to pinch)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Other skin GVHD features (NOT scored by BSA)

Check all that applies:

☐ Hyperpigmentation
☐ Hypopigmentation
☐ Poikiloderma
☐ Severe or generalized pruritus
☐ Hair involvement
☐ Nail involvement
☐ Abnormality present but explained entirely by non-GVHD documented cause (specify): __________________________

• How is progressive hyperpigmentation documented? document as cGVHD if present outsides areas of sun exposure and positive histology

• Why are hypopigmentation and poikiloderma not counted: not reversible – does not mean that it is not cGVHD, new poikiloderma should be counted but can not be used for response assessment
Clinical challenges in NIH grading – skin grading

- How is inactive cGVHD graded if irreversible sclerosis is present? – *document as it is with comment – withdrawal of IS possible?*
- Is an isolated maculopapular erythema regarded as cGVHD if onset at day 200 and histology reveals lichenoid cGVHD - *regarded as cGVHD?*
- How to differentiate between fasciitis and skin involvement especially of the lower legs? - *Primary involvement of cutaneous adnexal structures (hair, sweat)? Any erythema? Signs of hyper- or depigmentation?*
- How to differentiate between superficial and deep sclerosis at the lower legs? - *often impossible due to lack of subcutaneous fat*
- How to document response in cutaneous sclerosis? – *PROM, IS, Size*
- Does a localized deep sclerosis justify diagnosis of severe cGVHD? – *Yes?*
Clinical challenges in NIH grading – oral diagnosis
Clinical challenges in NIH grading – oral diagnosis and grading

- Hyperkeratotic changes are also present in healthy individuals and may be a premalignant lesion—consider biopsy if treatment decision depends.
- Sicca symptoms are not well captured—Dry bread possible? Score sicca according to extent of impairment (mild versus moderate).
- Mucoceles fluctuates over time—removed from scoring.

Scoring:
- Mild cGVHD: if signs are present but do not interact with oral intake.
- Moderate cGVHD: if signs are present and partial limit oral intake (hot meals, drinking required during eating, toothpaste?)
- Severe cGVHD: major impairment of oral intake (weight loss, only fluid meals, pain medication?)
Clinical challenges in NIH grading – ocular diagnosis and grading

- Current staging does not require ophthalmologist but does not cover Blepharitis nor intensity of inflammation
- Schirmer’s Test had been removed from diagnostic criteria
- Patients complaining about eye abnormalities post Tx should be referred to an ophthalmologist
- Baseline assessment on day 100 suggested
- Grading scale for ophthalmologist available (Dietrich et al Cornea 2011)

Grade according to the extent of impairment:
- Mild – sicca but no need for frequent lubricant use (<3 times per day) and no impairment of vision nor need for special eyeware
- Moderate – lubricant use > 3 times per day required
- Severe: impaired vision, special eyeware required, unable to work
Clinical challenges in NIH grading – ocular grading

Patient using special contact lenses for treatment of ocular GVHD need to be graded as severe despite symptomatic relief requirement of eye drops 2 only times a day – (Severe dry eye symptoms significantly affecting ADL or special eyeware to relieve pain)

➢ The same applies for punctal plugs for moderate
Clinical challenges in NIH grading – liver diagnosis and grading

• Diagnosis is made clinically – rarely histologically despite numerous differential diagnoses
• Change of treatment was reported in 36 – 67% of patients biopsied
• Indication for liver biopsy (Stift, J., Longerich, T. et al Virchows Arch 2013):

**Recommendations for performing hepatic biopsy:**
Liver biopsy should be limited to those patients, for whom the expected histological and/or microbiological findings are expected to have potential consequences for their therapeutic management. Liver biopsy is not required in clinical practice, if clinical symptoms are compatible with GvHD (especially if other organ manifestations are apparent) and respond to therapy. If the bleeding risk is acceptable, a liver biopsy is indicated in alloHCT patients with abnormal liver function who do not respond to seemingly adequate treatment and after non-invasive diagnostic steps have failed to establish a definitive diagnosis.
Clinical challenges in NIH grading – liver grading

• Severity was graded according to ALT, AST, AP and Bili until 2014

• Current proposal: only ALT, AP and Bilirubin is applied

<table>
<thead>
<tr>
<th>LIVER</th>
<th>Normal total bilirubin, ALT and AP &lt; 3 x NUL</th>
<th>Normal total bilirubin and ALT or / and AP ≥3 x NUL &lt; 5x NUP</th>
<th>Elevated total Bilirubin but ≤3 mg/dL NUL, ALT &gt; 5 x NUL</th>
<th>Elevated total bilirubin &gt; 3 x NUL</th>
</tr>
</thead>
</table>

☐ Abnormality present but explained entirely by non-GVHD documented cause (specify): _______________________________

Currently expert consensus – requires validation

Analysis of the data of the BUM trial revealed significant downgrading between mild and no GVHD: old definition: none (n=65), mild (n=67)

new definition: none (n=150), mild (n=11)
### Clinical challenges in NIH grading – lung diagnosis and grading

#### LUNGS**

<table>
<thead>
<tr>
<th>Symptoms score:</th>
<th>No symptoms</th>
<th>Mild symptoms (shortness of breath after climbing one flight of steps)</th>
<th>Moderate symptoms (shortness of breath after walking on flat ground)</th>
<th>Severe symptoms (shortness of breath at rest; requiring 0₂)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung obstructive function score:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV₁</td>
<td>FEV₁ ≥ 80%</td>
<td>FEV₁ 60-79%</td>
<td>FEV₁ 40-59%</td>
<td>FEV₁ ≤ 39%</td>
</tr>
</tbody>
</table>

Pulmonary function tests

- Not performed

Abnormality present but explained entirely by non-GVHD documented cause (specify):

- Screening of asymptomatic patients required
- For classical BO only FEV₁ applied for severity grading and response assessment
- If interstitial changes are involved complete LFT incl. DLCO required
Clinical challenges in NIH grading – diagnostic criteria

Diagnostic criteria:

- FEV1/VC < 0.7
- FEV1 decline of 10% over less than 2 years (non reversible) (baseline required)
- Absence of infection (assessment after resolution of infection)

Supporting criteria:

- Presence of other signs of cGVHD
- Signs of BO in the chest CT scan

Problems:

- BO may also be associated with additional restrictive changes
- Late IPS /lymphocytic alveolitis, BOOP, restriction of chest wall are not covered by diagnostic criteria for BO
Inspiration

FEV1 = 57%
FEV1/VCmax = 108%
MEF 25 = 58%
RV%TLC 178%

DLCO 65%
RV 152%
Expiration

- Detection of BO requires CT scan in Expiration
MinIP in Expiration
Clinical challenges in NIH grading – genital diagnosis and grading

• Grading requires gynecological work up
• Routine screening required (especially postmenopausal women do not report symptoms spontaneously, incidence in screening cohorts much higher than spontaneously reported)

Female genitalia – severity grading:
1) **Mild**: erythema on vulvar mucosal surfaces, vulvar lichen-planus or vulvar lichen-sclerosis
2) **Moderate**: erosive inflammatory changes of the vulvar mucosa, fissures in vulvar folds
3) **Severe**: labial fusion, clitoral hood agglutination, fibrinous vaginal adhesions, circumferential fibrous vaginal banding, vaginal shortening, synechia, dense sclerotic changes, and complete vaginal stenosis
Clinical challenges in NIH grading – genital diagnosis and grading

• For the first time male genital involvement can be graded
• Diagnostic features are lichen planus-like or lichen sclerosis-like features and phymosis or urethral scarring or stenosis.

Male genitalia - severity grading:

- Mild – lichen planus-like feature
- Moderate – lichen sclerosis-like feature or moderate erythema
- Severe – phymosis or urethral/meatal scarring

<table>
<thead>
<tr>
<th>GENITAL TRACT (See Supplemental table†)</th>
<th>□ No signs</th>
<th>□ Mild signs‡ and females with or without discomfort on exam</th>
<th>□ Moderate signs‡ and may have signs* of discomfort on exam</th>
<th>□ Severe signs‡ with or without symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Check all that applies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>□ Not examined</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currently sexually active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>□ Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>□ No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

□ Abnormality present but explained entirely by non-GVHD documented cause (specify): ____________________________

D. Wolff / Dept. of Medicine III, University of Regensburg
Clinical challenges in NIH grading – fascia diagnosis and grading

- Diagnosis is made clinically
- Typical early signs are pale erythema and edema of the fascia with tenderness but no dermal or epidermal exanthema or erythema
- Late signs are fibrosis of the fascia, lymphedema, decreased range of motion
- Signs of nerve or muscle compression
### JOINTS AND FASCIA

**P-ROM score** (see below)  
Shoulder (1-7): ___  
Elbow (1-7): ____  
Wrist/finger (1-7): ___  
Ankle (1-4): ___

<table>
<thead>
<tr>
<th>□ No symptoms</th>
<th>□ Mild tightness of arms or legs, normal or mild decreased range of motion (ROM) AND not affecting ADL</th>
<th>□ Tightness of arms or legs OR joint contractures, erythema thought due to fasciitis, moderate decrease ROM AND mild to moderate limitation of ADL</th>
<th>□ Contractures WITH significant decrease of ROM AND significant limitation of ADL (unable to tie shoes, button shirts, dress self etc.)</th>
</tr>
</thead>
</table>

□ Abnormality present but explained entirely by non-GVHD documented cause (specify): ___
Clinical challenges in NIH grading – GI diagnosis and grading

<table>
<thead>
<tr>
<th>GI Tract</th>
<th>No symptoms</th>
<th>Symptoms without significant weight loss* (&lt;5%)</th>
<th>Symptoms associated with mild to moderate weight loss* (5-15%) or moderate diarrhea without significant interference of daily living</th>
<th>Symptoms associated with significant weight loss* (&gt;15%), requires nutritional supplement for most calorie needs or esophageal dilation or severe diarrhea with significant interference of daily living</th>
</tr>
</thead>
<tbody>
<tr>
<td>Check all that applies:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>☐ Esophageal web/proximal stricture or ring</td>
<td>☐ Dysphagia</td>
<td>☐ Anorexia</td>
<td>☐ Nausea</td>
<td>☐ Vomiting</td>
</tr>
<tr>
<td>☐ Abnormality present but explained entirely by non-GVHD documented cause (specify): ________________________________</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

➤ Symptoms of acute GVHD are now captured
Conclusions

✓ NIH grading improves documentation and prevents ignorance of organ manifestations and permits response assessment
✓ Indication for histological confirmation defined
✓ Clarification and adjustments of definitions have been added in the 2014 consensus
➢ Validation of impact in clinical routine required
➢ Critical data for clinical routine versus trials are to defined
➢ Dissemination of knowledge required
➢ Phase IIb and Phase III trials evaluating existing treatment options required
➢ Diagnostic criteria (role of histopathology, biomarker) need to be validated
Future Meetings of the German-Austrian-Swiss-GVHD-consortium

Workshop on Cutaneous GVHD, Hamburg
7. November 2014

Workshop on Paediatric and Neurological GVHD
Wiesbaden
8. May 2015

Workshop on gastrointestinal GVHD Regensburg

www.gvhd.eu