CrossMark

ORIGINAL ARTICLE

Consensus diagnostic histopathological criteria for acute gastrointestinal graft versus host disease improve interobserver reproducibility

Andreas Kreft[1] · Anja Mottok[2] · Ildiko Mesteri[3] · Diana M. Cardona[4] · Anne Janin[5,6] ·
Anja A. Kühl[7] · Mindaugas Andrulis[8] · Andrea Brunner[9] · Howard M. Shulman[10] ·
Giovanni Negri[11] · Alexandar Tzankov[12] · Elisabeth Huber[13] · on behalf of the
Gastrointestinal Pathology Group of the German-Austrian-Swiss GvHD Consortium

Abstract Graft versus host disease (GvHD) is a clinically important complication after allogeneic hematopoietic stem cell transplantation (HSCT). Its diagnosis relies on clinical and histopathological findings. In order to evaluate and improve inter-institutional diagnostic agreement on histological diagnosis and grading of acute gastrointestinal GvHD, we conducted a round robin test, which included 33 biopsies from 23 patients after HSCT. Five pathologists from different institutions independently evaluated the original sections from the biopsies submitted for diagnosis. Based on their results, consensus qualitative criteria for the assessment of typical histological features of GvHD (e.g., apoptosis, crypt destruction, mucosa denudation) were proposed, including detailed descriptions as well as histological images. In a second round robin test with involvement of the same pathologists, the reproducibility of both diagnosis and grading had improved. Remaining differences were mostly related to differential diagnostic considerations, including viral infection or toxic side effects of medication, which should be resolved by integrating histopathological findings with proper clinical information.

Keywords Intestinal GvHD · Diagnosis · Histology · Consensus criteria

Introduction

Acute intestinal graft versus host disease (GvHD) is a severe potentially life-threatening complication after allogeneic bone marrow or peripheral hematopoietic stem cell transplantation (HSCT). Central to its pathogenesis are engrafted donor immune cells causing an allo-immunoreaction that leads to epi-

✉ Andreas Kreft
andreas.kreft@unimedizin-mainz.de

1 Institute of Pathology, University Medical Center Mainz,
Langenbeckstr. 1, 55101 Mainz, Germany

2 Institute of Pathology, University of Würzburg and Comprehensive
Cancer Centre Mainfranken, Josef-Schneider-Str. 2,
97080 Würzburg, Germany

3 Institute of Pathology, Medical University of Vienna, AKH,
Währinger Gürtel 18 –20, A-1090 Vienna, Austria

4 Department of Pathology, Duke University Medical Center, BOX
3712, Durham, NC 27713, USA

5 Université Paris Diderot, Laboratoire de Pathologie, UMR_S1165,
F-75010 Paris, France

6 Department of Pathology, AP-HP, Hôpital Saint-Louis,
F-75010 Paris, France

7 Department of Medicine I for Gastroenterology, Infectious Disease
and Rheumatology, Campus Benjamin Franklin, Charité -
Universitätsmedizin Berlin, Hindenburgdamm 30,
12203 Berlin, Germany

8 Institute of Pathology, Heidelberg University Hospital,
Heidelberg, Germany

9 Department of Pathology, Division of General Pathology, Medical
University of Innsbruck, Muellerstrasse 44,
A-6020 Innsbruck, Austria

10 Seattle Cancer Center Alliance, Room G7-910, 825 Eastlake ave E,
Seattle, WA 98109, USA

11 Department of Pathology, Central Hospital Bolzano, Via Boehler 5,
I-39100 Bolzano, Italy

12 Institute of Pathology, University Hospital Basel, Schoenbeinstr. 40,
CH-4031 Basel, Switzerland

13 Institute of Pathology, University Hospital Regensburg,
Franz-Josef-Strauss-Allee 11, 93053 Regensburg, Germany

⸖ Springer

thelial cell apoptosis, inflammation, and tissue injury [17]. The diagnosis of GvHD relies on clinical symptoms and—especially in non-diagnostic or precarious clinical settings—on histopathological findings [15]. However, assessment of severity of GvHD on histology, by grading the characteristic histological features, is often only weakly correlated with the clinical course of the disease [1, 16, 21, 23].

Difficulties in histopathological diagnosis of GvHD may result from different causes. In various proposed diagnostic scoring systems mostly classical, fully developed features are considered [10, 12, 14, 20]. However, in early cases with subtle features as well as in cases with heterogeneous features, subjective interpretation by individual pathologists may result in interindividual differences in both diagnosis and grading. Discrepancies may be marginal when pathologists from single or cooperating institutions have an opportunity to adjust each other's diagnostic criteria [5, 20], but when there is no such exchange of experience, they could more profoundly affect inter-institutional reproducibility [19, 24]. Additionally, other diagnostic considerations include infection [2, 3] and toxic side effects of applied therapies [11, 18], which are both frequent in patients after HCT and could mimic GvHD histologically as well as clinically.

In order to evaluate the interobserver variability of the histological diagnosis of acute intestinal GvHD, we conducted an international round robin test with the original slides used for histopathological diagnosis. Based on the obtained results, we proposed consensus criteria for morphological diagnosis. This consensus includes subtle features of early disease and of gray zones between different histologic grades, as well as classical characteristics of advanced disease, all supported by photographic images.

Materials and methods

Materials for study

The round robin test on acute gastrointestinal GvHD, in which 5 pathologists participated, included 33 biopsies from 23 patients after allogenic HSCT. Since an initial evaluation of scanned whole slides was found to be insufficient in assessing delicate histological details needed for diagnosis, e.g., apoptotic bodies, the slides used for the original histological diagnosis were sent around for evaluation. We included 23 colonic, 5 small intestinal, and 4 gastric biopsies (Table 1) (one colonic and one esophageal biopsy were also sent, but not included in the presented evaluation because of a short interval after HSCT of only 8 days for the former and of having only one biopsy at that site for the latter). According to current diagnostic standards of the participating institutions, the evaluated slides were stained with hematoxylin-eosin (H&E) and additional conventional as well as immunohistological stains if

deemed appropriate. Each H&E stained slide contained up to 8 tissue sections of a single biopsy. Additional stains included periodic acid-Schiff (PAS) stains for 6 biopsies, Giemsa stains for 4 biopsies, Epstein-Barr virus (EBV)-encoded RNA (EBER) in situ hybridization for 2 biopsies as well as immunostains (2 biopsies stained for cytomegalovirus (CMV), 2 for herpes simplex virus type one and two, and for CD20, CD56, CD138, MUM1, kappa, and lambda light chain on a single biopsy). Caspase immunostaining was not provided, since none of the institutions used it for routine GvHD diagnosis.

The cases selected for review were contributed by institutes of pathology in Switzerland (Basel (A.T.)), Italy (Bolzano (G.N.)), and Germany (Regensburg (E.H.), Mainz (A.K), and Würzburg (A.M.)). All these institutions are associated with clinics which perform allogeneic HSCT. Cases for this study were not obtained consecutively or randomly, but rather, selected biopsies were specifically chosen to represent a spectrum of cases considered classical, difficult, rare, borderline, or not diagnostic for GvHD. The sample therefore included the full gamut of all histologic grades of acute gastrointestinal GvHD.

Methods for histologic evaluation

The participants were requested to provide a diagnosis (positive or negative for GvHD, or an alternative diagnosis), a Lerner score in case of GvHD and a semiquantitative evaluation of the number of apoptotic bodies, crypt destruction, and mucosa denudation, as outlined in Table 2 [8, 10]. The Lerner score is the most widely used score for the diagnosis of acute GvHD in Germany, Austria, and Switzerland.

The accompanying clinical information initially provided on the request form by endoscopists or hematologists was made available to the participants, in order to simulate a standard histopathological diagnostic setting [15]. Lists the submitting institution, relevant patient data as available on the surgical biopsy request form, the type of tissue(s), and the GvHD histologic Lerner score(s). As we focused on histology, follow-up data or correlation with the clinical course was not included in the present study.

The findings of the initial round robin review were discussed in Basel (Switzerland) in 2012 at a meeting of the German-Austrian-Swiss GvHD Consortium. Differences between participants in assessment of histological criteria of GvHD (i.e., apoptosis, crypt destruction, and mucosa denudation) and their interpretation were identified. The need for consensus on criteria, including those covering advanced as well as subtle morphological changes, was recognized.

A questionnaire, containing questions concerning diagnostic strategies and histological criteria, was developed and sent to the 5 pathologists who took part in the initial round robin test and 7 others from institutions in which HSCT is performed (A.J. from France, M.A. and A.A.K. from Germany,

Table 1 Patient characteristics

Biopsy data

Sending institution	Patient						Assigned Lerner grade[a]	
	Age	Gender	Disease	Number of biopsies	Days after HSCT	Type of tissue	First round	Second round
Basel	48	M	B-ALL	1	10	Stomach	2–3	0–3
	62	M	Myeloma	1	113	Colon	0–1	1
	25	M	B-ALL	1	48	SI	3–4	3–4
	47	M	Myeloma	1	26	SI	0–3	0–1
Bolzano	60	M	AML	2	34	Colon	0–2	0–1
						Colon	1–2	1
	49	F	AML	1	75	Colon	1–3	1–2
	55	M	AML	3	120	Colon	3–4	3–4
						Colon	3–4	3–4
						Colon	2–4	1–4
Mainz	49	M	Hodgkin	2	155	Colon	0–1	1
						Colon	1	1
	70	M	MDS	1	77	Colon	0–2	0–1
	70	F	AML	1	56	Colon	0–1	0–1
	30	F	ALL	4	121	SI	2–3	2–3
						Stomach	0–2	0–1
						Colon	3–4	2–4
						Colon	3–4	2–4
Regensburg	60	M	AML	1	34	Colon	0–1	0–1
	49	F	AML	3	75	SI	0–3	0–4
						Stomach	2–3	2–3
						Stomach	1–2	1–3
	55	M	AML	1	120	Colon	0	0
	40	F	AML	1	21	SI	1–2	0–3
Würzburg	48	M	Myeloma	1	749	Colon	1	1–4
	72	M	AML	1	112	Colon	1–2	1–2
	34	F	AML	2	220	Colon	2–4	4
						Colon	1–4	4
	62	F	MDS	1	23	Colon	1–2	2
	62	F	MDS	1	26	Colon	1–4	3–4
	51	F	AML	1	24	Colon	1	1
	62	F	CLL	1	224	Colon	1–2	1
	56	M	Myeloma	1	113	Colon	1	1

ALL acute lymphoblastic leukemia, *MDS* myelodysplastic syndrome, *AML* acute myeloid leukemia, *CML* chronic myeloid leukemia, *CLL* chronic lymphocytic leukemia, *SI* small intestine

[a] The range Lerner grades of both first and second round of the robin test given by the participants are provided

I.M. and A.B.. from Austria, and D.C. and H.S. from USA). Digital photomicrographs of features of GvHD in gut biopsies from patients after HSCT with potential diagnostic significance were circulated between participants. The results were discussed at a follow-up meeting of the Consortium in Mainz (Germany) in 2013. A written consensus, including diagnostic strategies, diagnostic criteria (Table 3), and representative photomicrographs (Figs. 1, 2, 3, 4, 5, and 6), was approved by all 12 participants. Subsequently, the slides used for the original histological diagnosis were again sent to the five institutions of the first robin test for reevaluation by applying the consensus criteria. The time lapse between the two evaluations was at least 1 year for each of the scoring pathologists.

Statistical evaluation was performed for interobserver agreement (in %) and for multiple observers using Fleiss generalized kappa. We found a high number of biopsies with

Table 2 Lerner classification of acute intestinal GvHD

Diagnostic category	Histological sign
Grade 0	Normal mucosa
Grade 1	Apoptosis of crypt or gland epithelium
Grade 2	Crypt or gland destruction
Grade 3	Focal mucosa denudation and/or ulceration
Grade 4	Diffuse mucosa denudation and/or ulceration

According to [8, 10]

diagnosis "GvHD" compared to "no GvHD" (high marginal heterogeneity), which could have resulted in a high chance-agreement probability with a relatively low kappa value [6], and therefore, we also computed a reliability coefficient with an adjusted change agreement, the AC1 of Gwet, which was shown to compensate for the effect of high marginal heterogeneity [22]. To compute the results, AgreeStat 2013.1 software (Advanced Analytics, LLC, Gaithersburg, MD, USA [7]) on Microsoft Excel 2013 (Microsoft, Seattle, WA, USA) was used.

Results

Histology

The histologic architecture was well preserved in most of the 24 biopsies of the colon. Occasionally, clusters of apoptotic cells were noted in crypt epithelium. The extent of crypt destruction varied and included severe damage, qualifying as crypt destruction. Some cases had features of crypt loss and/or mucosa denudation. CMV immunohistochemistry was done in one biopsy from the colon but was negative. Of the 5 biopsies of the small intestine, 3 were from the duodenum, and of 2 others, complete mucosal denudation hampered histological identification of their origin. The duodenal biopsies showed well preserved architecture, and a variable number of apoptotic bodies, with occasional crypt destruction or mucosal denudation. In one biopsy, an atypical plasmacytoid infiltrate was suggestive of post-transplantation lymphoproliferative disease (PTLD). This case was also positive for CMV and EBER. The four gastric biopsies showed apoptotic cells and some crypt destruction, up to mucosal denudation in two samples.

Consensus criteria

The authors agreed that for a diagnosis GvHD or its grading, the most advanced tissue alteration in a given biopsy needed to be considered. The consensus diagnostic criteria for apoptosis included shrinkage of crypt epithelial cells in combination with condensed nuclear chromatin and cytoplasmic eosinophilia (Fig. 1). As an alternative, a minimum of two particles of nuclear debris in one area or signs of phagocytosis of cell debris were accepted. In contrast, close contact of intraepithelial lymphocytes with apoptotic epithelial cells (satellitosis), even though characteristic of apoptosis, was not considered necessary for establishing apoptosis. Apoptotic cells in luminal surface epithelium or in the lamina propria were not considered diagnostic for GvHD, but part of the physiological cell turn over [4]. Isolated cells with condensation of nuclear chromatin, shrinkage, or

Table 3 Consensus on diagnostic criteria of intestinal acute GvHD

|	Apoptosis	Crypt destruction	Mucosa denudation
Diagnostic criteria	• Cell shrinkage in combination with increased eosinophilia of the cytoplasm and condensed nuclear chromatin • Apoptotic bodies/karyorrhectic debris (at least 2—over an area of the size of an epithelial cell) • Phagocytosis of apoptotic bodies or cells • For the diagnosis of intestinal aGvHD, only apoptosis in epithelial cell layers of the crypts should be considered	• Apoptosis (destruction) of at least one third of the crypt or gland epithelial cells. The involved crypt should have at least half of the diameter of a normal crypt • Widening of gland or crypt lumen (exploding crypt) and/or flattening of most of the epithelium, filled with cell detritus and/or inflammatory cells	• No visible epithelial cells on the surface of the mucosa and fibrin deposits on the surface. Focal accumulation of inflammatory cells and/or widening of the superficial blood vessels would support the diagnosis
Changes not sufficient for diagnosis, if isolated	• Condensation of nuclear chromatin • Cell shrinkage • Cell ballooning	• Focal atrophy of the crypt or gland epithelial cells • Crypt abscess • Loss of crypts or glands may reflect previous GvHD as well as other causes of mucosa injury. It is not a sign of active GvHD	• No visible epithelial cells on the (luminal) surface of the mucosa, without fibrin deposits or inflammatory or vascular reaction

According to [8, 10]
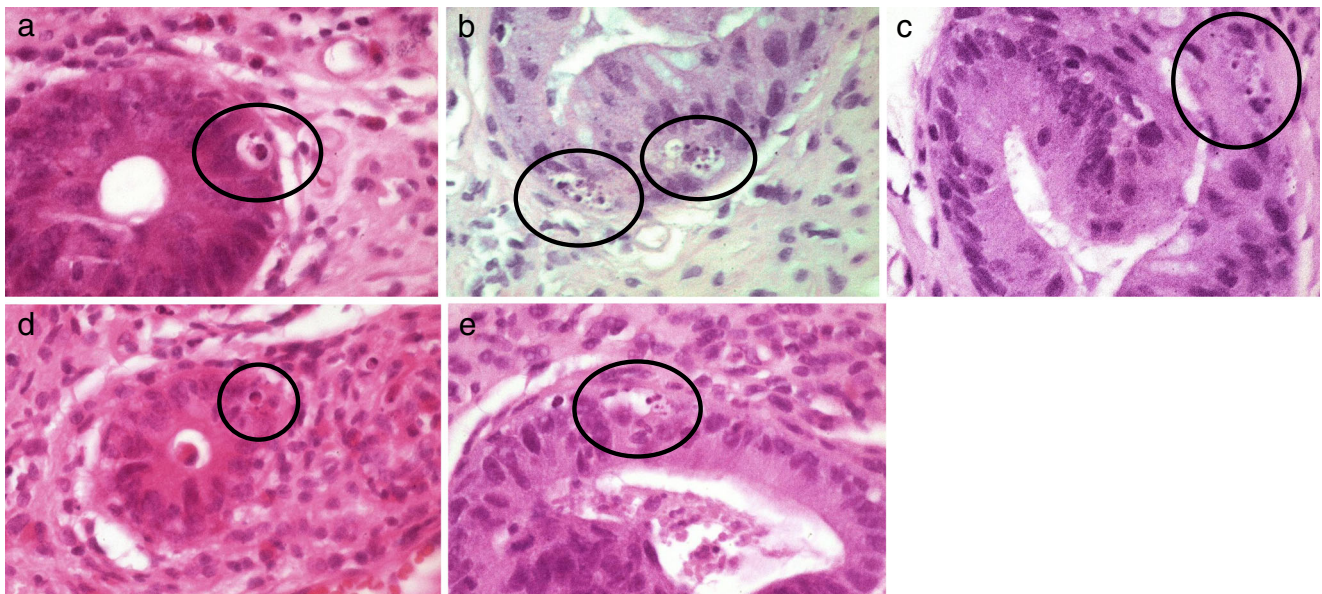
Fig. 1 Typical examples of crypt cell apoptosis (HE, original magnification ×400): **a**, **d** each two apoptotic bodies (only one would not have been diagnostic), with surrounding halo; **b**, **c**, and **e** more apoptotic bodies, easy to recognize, with a "dusty" appearance
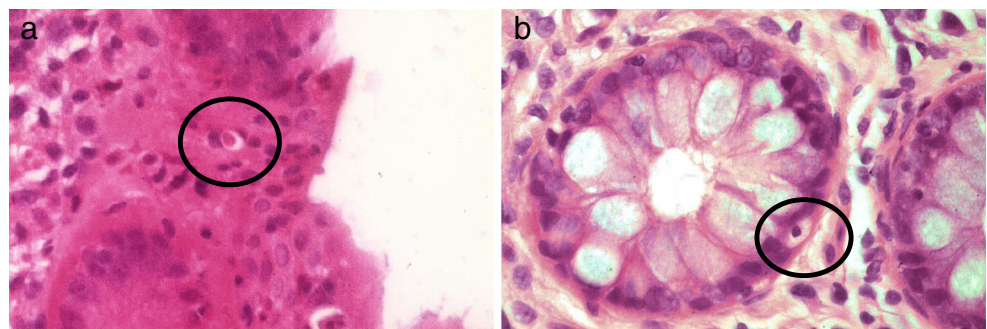
ballooning were also not considered sufficient for establishing apoptosis, although they do indicate cell damage (Fig. 2).

As criteria for crypt- or gland destruction, destruction of at least one third of the enterocytes comprising the circumference of a cross-sectioned crypt was deemed necessary. Dilation of a crypt accompanied by epithelial flattening and luminal cell debris was considered characteristic of an apoptotic crypt or gland abscess (Fig. 3). Without these signs of damage of the epithelial cells or in the presence of a lesser degree of atrophy of epithelial cells, crypt destruction could not be concluded (Fig. 4). The mere loss of crypts might be a remnant of previous episodes of GvHD, in correlation with clinical symptoms and therapy refractoriness [12, 15], but not a sign of acute GvHD.

Lack of epithelial cells on the mucosal surface was considered essential to conclude mucosal denudation. In addition, deposition of fibrin or accumulation of inflammatory cells and/or widening of small blood vessels was required to avoid confusion with artificial damage (Figs. 5 and 6).

A complete list of histological consensus criteria is given in Table 3.

Interobserver agreement

First round robin test

Before the consensus, all 5 observers agreed in 24 of the 33 cases on the diagnosis of GvHD (positive 23, negative 1) (Table 1). In 4 of the remaining 9 cases, there was one divergent interpretation, while in 5 more cases, 2 pathologists rendered a divergent diagnosis. The discordances were related to concerns of focal or subtle changes (4 cases), the possibility of concurrent infections (CMV in one case, EBV in one case with in addition PTLD), of therapy-related toxic changes (2 cases), or of advanced tissue damage without determinable cause (1 case). The agreement was 84 % (kappa 0.347, Gwet's AC1 0.792) for the diagnosis of GvHD.

A divergence in grading of acute GvHD occurred in 20 of the 24 cases diagnosed as GvHD by all participants. Differences of one grade occurred in 16, of 2 grades in 2, and of 3 grades also in 2 cases. Divergence in 7 of these cases related to possible CMV infection and in one to therapy-related toxic changes. The agreement for Lerner grading was 48 % (kappa
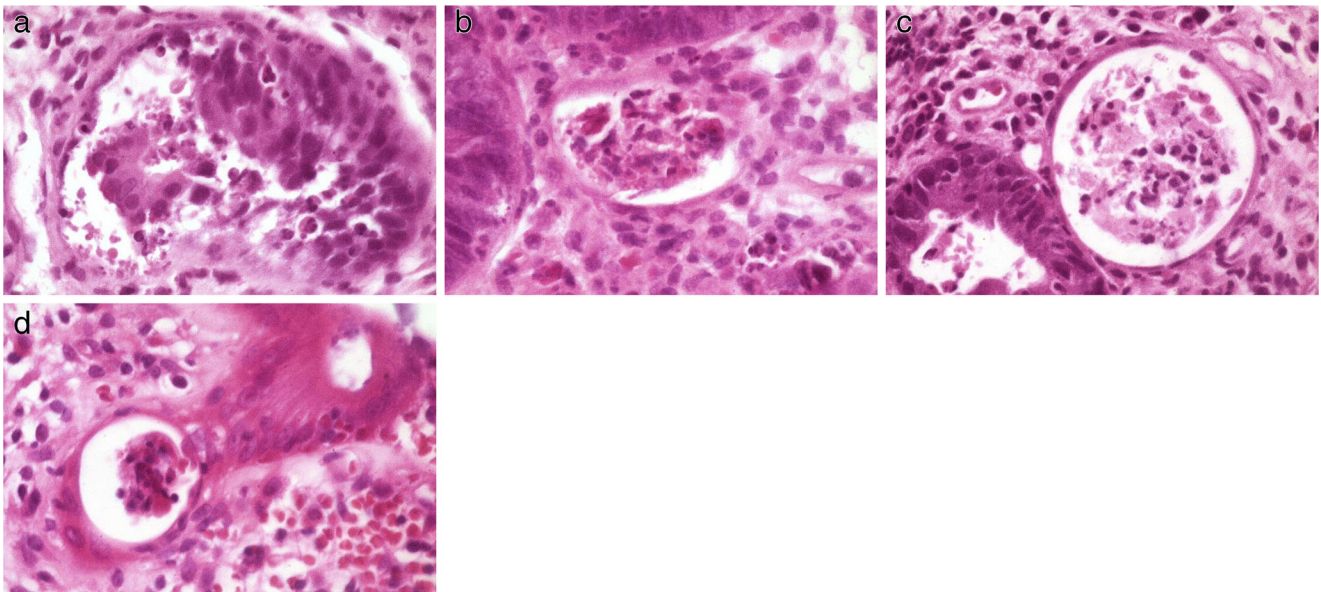

Fig. 2 These features were not regarded to be diagnostic for apoptosis (HE, original magnification ×1000): **a** revealed only a single eosinophilic element of uncertain provenience, **b** possibly a lymphocyte

Fig. 3 Typical examples of crypt destruction (HE, original magnification ×400): **a** with destruction of more than 1/3 of the circumferential crypt epithelium by confluent apoptosis; **b** and **c** with flattened epithelium; and **d** the connection to normal appearing superficial parts of the crypt

0.322; Gwet's AC1, 0.323). When the grading categories were taken together as low grade (Lerner grade 1 and 2) or high grade (grade 3 and 4), the agreement increased to 68 % (kappa 0.457; Gwet's AC1, 0.552).

Second round robin test

After the written consensus, an agreement on the diagnosis of GvHD was reached in 25 of the 33 cases (positive 23, negative 2). Of the remaining 8 divergent cases, 2 had one divergent diagnosis while the remaining 6 had 2 divergent diagnoses.

Divergent diagnoses were due to focal or subtle disease (2 cases), the possibility of concurrent infections (CMV in one case, EBV in one case with in addition PTLD), of therapy-related toxic changes (2 cases), or of advanced tissue damage without determinable cause (1 case). The agreement for a diagnosis of GvHD increased only marginally to 85 % (kappa 0.396; Gwet's AC1, 0.805).

As to grading of the 25 cases diagnosed as GvHD, 10 observers agreed on the grade but 13 diverged with differences of one grade in 8, of 2 grades in 3, and of 3 grades in 2 cases. In 6 of these cases, a possibility of CMV coinfection and, in one case, toxic changes were considered. The
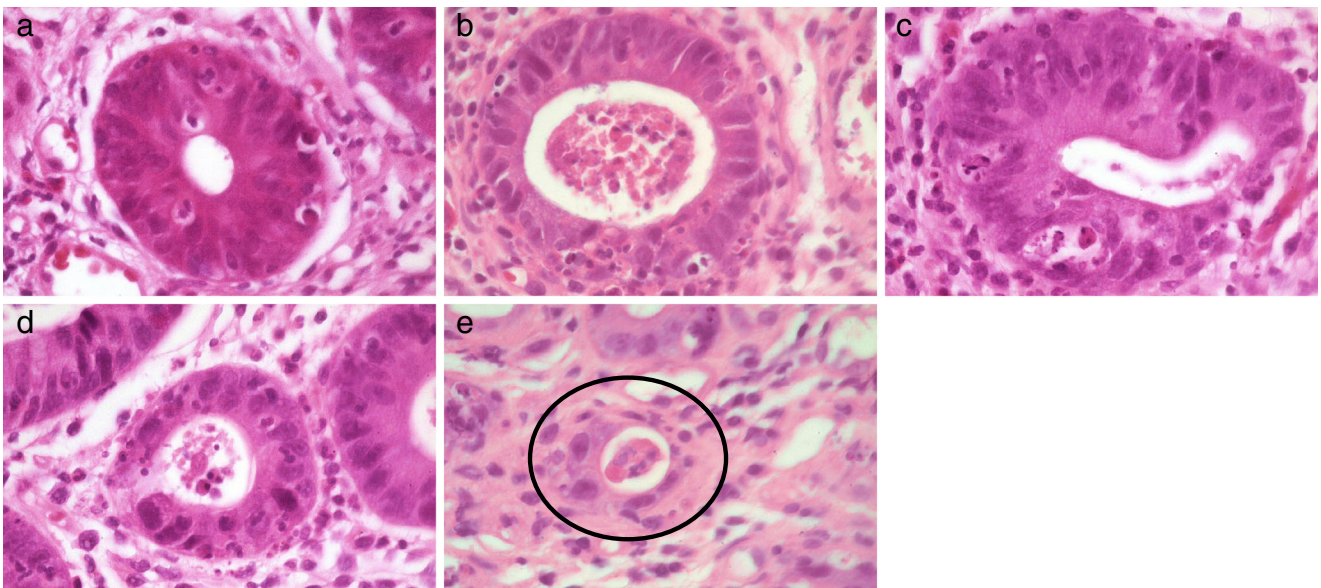

Fig. 4 Crypts not considered representing crypt destruction (HE, original magnification ×400): Crypt epithelium is not flattened; **a** reveals some inflammatory cells; in **b**, **c**, and **d,** apoptosis is present in a small portion of the epithelial cells; and **e** a small crypt filled with detritus and focal atrophy of the epithelium
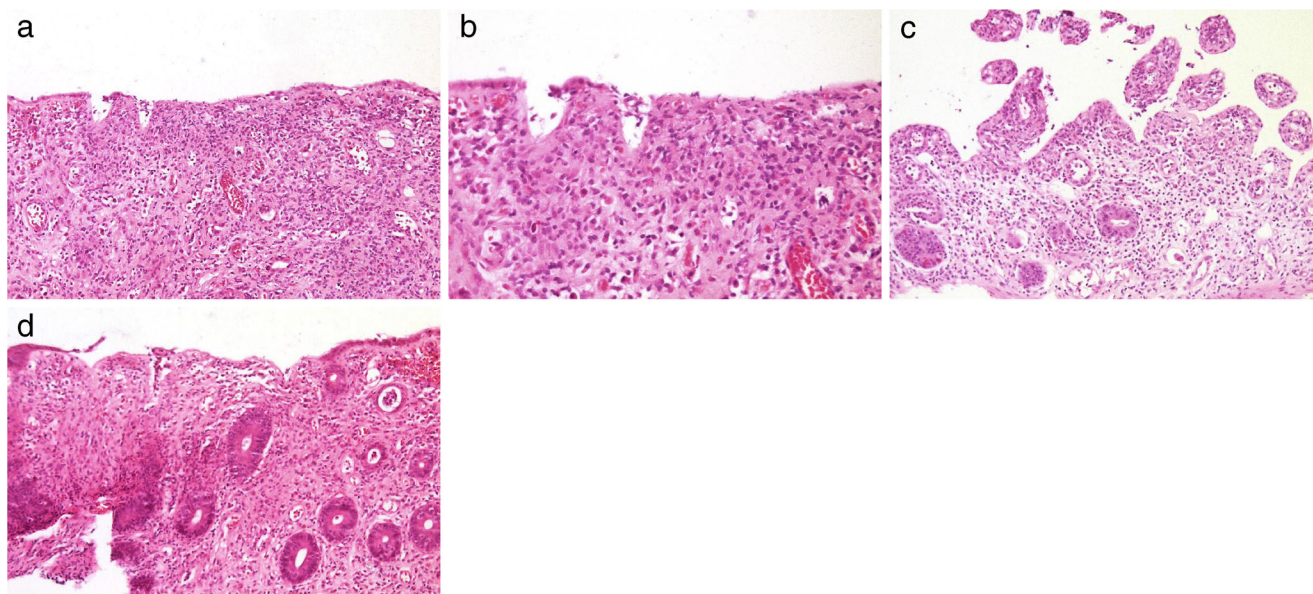
Fig. 5 Typical examples of mucosa denudation (HE, original magnification **a**, **c**, **d** ×100, **b** ×400; with **b** showing a detail of **a** at higher magnification.) These examples show the loss of the surface epithelial layer and of many crypts as well as some inflammatory reaction of cells and fibrin at the surface

interobserver agreement for grading increased to 60 % (kappa 0.455; Gwet's AC1, 0.512), and by clustering the grading categories to low grade and high grade, the agreement increased to 74 % (kappa 0.55; Gwet's AC1, 0.634).

Discussion

Shortcomings of established grading schemes such as Lerner histologic grading [10] and related modifications [14, 20] have been pointed out in earlier [15] and recent [16] consensus documents dealing with the histopathological diagnosis of gastrointestinal GvHD. Firstly, diagnostic criteria for a minimal degree of GvHD are not standardized and subject to individual variation. Secondly, tissue damage evolves, and during the course of disease, different steps may cumulate. When a biopsy is taken at the onset of gut GvHD, the diagnosis may be made when only a few apoptotic cells are found, without marked mucosal crypt damage or loss. In our study, Lerner

grading included threshold signs as well as cell damage accumulated in later stages of the disease.

The initial histological evaluation resulted in substantial disagreement on diagnosis and grading of GvHD. This was partly due to a lack of precision in defining apoptosis, crypt destruction, and mucosal denudation. With the additional input from a larger panel of reviewers, we developed consensus definitions of essential histological features which constitute the basis for a diagnosis of gut GvHD. With these consensus diagnostic criteria, a measurable improvement in the reproducibility of diagnosis (negative or positive for GvHD) and grading (particular of low and high grade GvHD) was attained. This is clinically relevant as early GvHD is not easily detectable by endoscopy [9], which leaves open other causes of diarrhea such as infection or therapy-related toxicity.

Consensus on criteria for a histological diagnosis of GvHD are particularly important when results from different institutions are compared, such as quantification—in order to establish a threshold for diagnosis—and grading [13]. Although correlation with the clinical course, which was not included in our study,
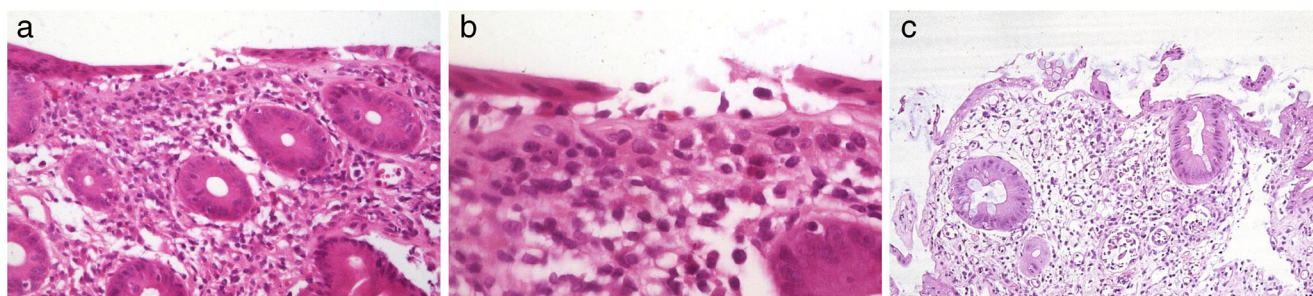

Fig. 6 Mucosal changes not accepted to be diagnostic for mucosa denudation (HE, original magnification **a** ×400, **b** ×1000, **c** ×250 with **b** showing a section of **a**. Few inflammatory cells without fibrin deposits, while the loss of the epithelium is most likely due to mechanic injury

would be necessary to validate diagnostic criteria, most participants considered up to two apoptotic cells per 4 mm^2 detected at a magnification of ×100 (using a ×10 eyepiece with a ×10 objective) as sufficient for a diagnosis of low-grade GvHD.

Consensus criteria also improved reproducibility of GvHD grading, from 48 to 60 %. Clustering of the four Lerner grades into low (grades 1 and 2) and high grade (grades 3 and 4) increased the agreement from 68 % (first round) to 74 % (second round). Using the Fleiss kappa for computing the interobserver agreement, we faced the "first kappa paradox" due to the high percentage of GvHD diagnoses of in our biopsies [6]. Kappa value was relatively low, especially for GvHD versus no GvHD. Gwet's AC1 values are less affected by prevalence and marginal probability and therefore provide a more stable reliability coefficient [7, 22]. We found Gwet's AC1 values to be closer to the percentage of agreement.

The remaining diagnostic divergences were mostly due to unresolved differential diagnoses, especially viral infection (such as CMV and adenovirus) and therapy-related toxicity, which may produce histological features similar to those of GvHD [11, 18]. This underscores the need for microbiological and immunohistological studies. It should be borne in mind that immunohistochemical evidence of a CMV infection does not exclude GVHD because GVHD, infection, and therapy-related toxicity are not mutually exclusive.

A potential limitation of our study is selection bias, as the included cases were neither prospectively nor randomly collected. Participants selected diagnostically difficult cases, which they considered of particular interest based on their experience, including biopsies not diagnostic for GvHD and various grades of acute GvHD. The kappa values for interobserver agreement should therefore not be regarded as a benchmark for a diagnostic standard, which would have to be established on a random sample or consecutive case series.

In summary, we have developed a set of criteria that define histological features of gut GvHD, notably apoptosis, crypt destruction, and mucosal loss. These can be applied to any subsequent study. When evaluating samples with minimal changes, perfect agreement is difficult to attain. We show that consensus diagnostic histological criteria improve interindividual reproducibility. Close interaction between pathologists and attending clinicians remains a cornerstone for adequate interpretation of histopathological observations, in providing arguments to eliminate critical differential diagnoses.

References

1. Abraham J, Janin A, Gornet JM, de Latour RP, Robin M, Xhaard A, de Fontebrune FS, Mary JY, Allez M, Porcher R, Socié G (2014) Clinical severity scores in gastrointestinal graft-versus-host disease. Transplantation 97:965–971
2. Castagnola E, Cappelli B, Erba D, Rabagliati A, Lanino E, Dini G (2004) Cytomegalovirus infection after bone marrow transplantation in children. Hum Immunol 65:416–422
3. Cho BS, Yahng SA, Kim JH, Yoon JH, Shin SH, Lee SE, Choi SM, Lee DG, Eom KS, Park G, Kim YJ, Kim HJ, Lee S, Min CK, Cho SG, Kim DW, Lee JW, Min WS, Park CW (2013) Impact of cytomegalovirus gastrointestinal disease on the clinical outcomes in patients with gastrointestinal graft-versus-host disease in the era of preemptive therapy. Ann Hematol 92:497–504
4. Dahl J, Greenson JK (2012) Colon. In: Mills SE (ed) Histology for pathologists, 4 edn. Wolter Kluver, Lippincott Williams & Wilkins, Philadelphia, pp 673–695
5. Epstein RJ, McDonald GB, Sale GE, Shulman HM, Thomas ED (1980) The diagnostic accuracy of the rectal biopsy in acute graft-versus-host disease: a prospective study of thirteen patients. Gastroenterology 78:764–771
6. Feinstein AR, Cicchetti DV (1990). High agreement but low Kappa: I. the problems of two paradoxes. J Clin Epidemiol 43: 543–549. doi: 10.1016/0895-4356(90)90158-L
7. Gwet KL (2012) Handbook of inter-rater reliability. Advanced analytics. LLC, Gaitersburg
8. Hymes SR, Alousi AM, Cowen EW (2012) Graft-versus-host disease: part I. Pathogenesis and clinical manifestations of graft-versus-host-ádisease. J Am Acad Dermatol 66:515
9. Kreisel W, Dahlberg M, Bertz H, Harder J, Potthoff K, Deibert P, Schmitt-Graeff A, Finke J (2011) Endoscopic diagnosis of acute intestinal GVHD following allogeneic hematopoietic SCT: a retrospective analysis in 175 patients. Bone Marrow Transplant 47:430–438
10. Lerner KG, Kao GF, Storb R, Buckner CD, Clift RA, Thomas ED (1974) Histopathology of graft-vs-host reaction in human recipients of marrow from HL-A-matched sibling donors. Transplant Proc 4: 367–371
11. Liapis G, Boletis J, Skalioti C, Bamias G, Tsimaratou K, Patsouris E, Delladetsima I (2013) Histological spectrum of mycophenolate mofetil-related colitis: association with apoptosis. Histopathology 63:649–658
12. Melson J, Jakate S, Fung H, Arai S, Keshavarzian A (2007) Crypt loss is a marker of clinical severity of acute gastrointestinal graft-versus-host disease. Am J Hematol 82:881–886
13. Nguyen C, Kastenberg D, Choudhary C, Katz L, DiMarino A, Palazzo J (2008) Is single-cell apoptosis sufficient for the diagnosis of graft-versus-host disease in the colon? Dig Dis Sci 53:747–756
14. Sale GE, McDonald GB, Shulman HM, Thomas ED (1979) Gastrointestinal graft-versus-host disease in man: a clinicopathological study of the rectal biopsy am. J Surg Pathol 4:291–299
15. Shulman HM, Kleiner D, Lee SJ, Morton T, Pavletic SZ, Farmer E, Moresi M, Greenson J, Janin A, Martin PJ, McDonald G, Flowers MED, Turner M, Atkinson J, Lefkowich J, Washington MK, Prieto VG, Kim SK, Argenyi Z, Diwan H, Rashid A, Hiatt K, Couriel D, Schultz K, Hymes S, Vogelsang GB (2006) Histopathologic diagnosis of chronic graft-versus-host disease: National Institut of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: II. Pathology working group report. Biology of Blood and Marrow Transplantation 12:31–47
16. Shulman HM, Cardona DM, Greenson JK, Hingorani S, Horn T, Huber E, Kreft A, Longerich T, Morton T, Myerson D, Prieto VG, Rosenberg A, Treister N, Washington K, Ziemer M, Pavletic SZ, Lee SJ, Flowers MED, Schultz KR, Jagasia M, Martin PJ, Vogelsang GB, Kleiner DE (2015) 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 21:589–603. doi:10.1016/j.bbmt.2014.12.031

17. Socié G, Blazar BR (2009) Acute graft-versus-host disease: from the bench to the bedside. Blood 114:4327–4336. doi:10.1182/blood-2009-06-204669

18. Soldini D, Gaspert A, Montani M, Reineke T, Rogler G, Odze R, Weber A (2014) Apoptotic enteropathy caused by antimetabolites and TNF-a antagonists. J Clin Pathol 67:582–586

19. Stift J, Baba H, Huber E, Federmann B, Fischer HP, Schmitt-Graeff A, Baurmann H, Bethge W, Schirmacher P, Wrba F, Greinix H, Fend F, Schwerdtfeger R, Shulman H, Wolff D, Longerich T (2014) Consensus on the histopathological evaluation of liver biopsies from patients following allogeneic hematopoietic cell transplantation. Virchows Arch 464:175–190

20. Washington K, Bentley RC, Green A, Olson J, Treem WR, Krigman HR (1997) Gastric graft-versus-host disease: a blinded histologic study am. J Surg Pathol 21:1037–1046

21. Weisdorf DJ, Hurd D, Carter S, Howe C, Jensen LA, Wagner J, Stablein D, Thompson J, Kernan NA (2003) Prospective grading of graft-versus-host disease after unrelated donor marrow transplantation: a grading algorithm versus blinded expert panel review. Biol Blood Marrow Transplant 9:512–518

22. Wongpakaran N, Wongpakaran T, Wedding D, Gwet K (2013) A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol 13:61

23. Zhou Y, Barnett MJ, Rivers JK (2000) Clinical significance of skin biopsies in the diagnosis and management of graft-vs-host disease in early postallogeneic bone marrow transplantation. Arch Dermatol 136:717–721

24. Ziemer M, Haeusermann P, Janin A, Massi D, Ziepert M, Wolff D, Greinix H, Hillen U (2013) Histopathological diagnosis of graft-versus-host disease of the skin—an interobserver comparison. J Eur Acad Dermatol and Venereol 28:1–10