The ISHLT chronic lung allograft dysfunction consensus criteria are applicable to pulmonary chronic graft-versus-host disease

Yifan Pang,[1] Ananth V. Charya,[2] Michael B. Keller,[3,4] Arlene Sirajuddin,[5] Yi-Ping Fu,[6] Noa G. Holtzman,[7] Steven Z. Pavletic,[7,*] and Sean Agbor-Enoh[4,8,*]

[1]National Heart, Lung, and Blood Institute, [2]Pulmonary Branch, National Heart, Lung, and Blood Institute, and [3]Critical Care Medicine Department, Clinical Center, National Institutes of Health, Bethesda, MD; [4]Department of Pulmonary and Critical Care Medicine, Johns Hopkins Hospital, Baltimore, MD; [5]Department of Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, MD; [6]Office of Biostatistics Research, National Heart, Lung, and Blood Institute, [7]Immune Deficiency Cellular Therapy Program, Center for Cancer Research, National Cancer Institute, and [8]Laboratory of Applied Precision Omics, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD

Key Points

- The adapted ISHLT criteria can identify patients with high-risk pulmonary cGVHD that may not be captured by the current NIH criteria.

- The adapted criteria have the potential to become a valuable tool to adequately phenotype and study pulmonary cGVHD.

Pulmonary chronic graft-versus-host disease (PcGVHD) is a devastating complication of allogeneic hematopoietic stem cell transplant (HCT). The 2014 National Institutes of Health cGVHD consensus criteria (NIH criteria) only captures bronchiolitis obliterans syndrome (BOS). In this study, we adapted the 2019 International Society for Heart and Lung Transplantation (ISHLT) criteria of chronic lung allograft dysfunction (CLAD) to define novel phenotypes of PcGVHD and compared the performance of this criteria with the NIH criteria to identify patients with high-risk PcGVHD. We reviewed consecutive patients in a cGVHD natural history protocol (#NCT00092235) and adapted the 2019 CLAD criteria (the adapted criteria) to define PcGVHD as post-HCT forced expiratory volume at 1 second < 80% predicted value, with 4 phenotypes: obstructive, restrictive, mixed obstructive/restrictive, and undefined. An independent adjudication committee evaluated subjects for diagnosis and phenotyping. We identified 166 (47.4%) patients who met the adapted criteria, including obstruction (n = 12, 3.4%), restriction (n = 67, 19.1%), mixed obstruction/restriction (n = 47, 13.4%), and undefined (n = 40, 11.4%). In these patients, less than half (n = 78) met the NIH criteria for BOS (NIH+); the rest (n = 88) did not (NIH−). The NIH− subjects showed increased risk of death compared with those without PcGVHD (hazard ratio = 1.88, 95% confidence interval = 1.20-2.95; P = .006) that was similar to NIH+ subjects (P = .678). Our study demonstrated the potential of the adapted criteria in identifying patients with high-risk PcGVHD that have been missed by the NIH criteria. The adapted criteria could become a valuable tool to better phenotype and study lung disease in cGVHD.

Introduction

Chronic graft-versus-host disease (cGVHD) is a complication of allogeneic hematopoietic stem cell transplant (HCT) that affects 30% to 70% of long-term survivors and indicates poor quality of life and increased nonrelapse mortality (NRM).[1,2] Pulmonary manifestations are one of the most difficult to treat conditions in

The full-text version of this article contains a data supplement.

cGVHD.[3] Bronchiolitis obliterans syndrome (BOS), characterized by small airway inflammation and fibrotic obliteration, has been definitively linked to cGVHD. The current diagnostic criteria for BOS, as established by the 2014 National Institutes of Health (NIH) cGVHD consensus criteria, require a new-onset irreversible obstructive ventilatory defect on pulmonary function testing (PFT) after HCT with a forced expiratory volume at 1 second (FEV1) < 75% of predicted normal values and FEV1/forced vital capacity (FVC) < 0.7, with or without evidence of air-trapping on expiratory thoracic computer tomography (CT), and absence of other causes of pulmonary dysfunction.[4] BOS occurs in 3.4% to 10% of HCT recipients and implicates a 2-year and 5-year NRM rate of approximately 50% and 80%, respectively.[5-7] The NIH criteria do not include nonobstructive phenotypes of pulmonary cGVHD (PcGVHD). However, restrictive pulmonary disease after HCT has been described, and the incidence of severe restrictive ventilatory defects (vital capacity < 60%) beyond 100 days after transplant is around 3%.[3,8] The clinical characteristics and associations of the nonobstructive pulmonary complications with HCT outcomes are not well elucidated.

Chronic lung allograft dysfunction (CLAD), an irreversible and progressive form of allograft failure in lung transplantation, manifests as a sustained decline in FEV1 by at least 20% from the subject's best PFTs. Based the on 2019 International Society for Heart and Lung Transplantation (ISHLT) CLAD Consensus Report, the definition of CLAD and CLAD phenotypes, including BOS, restrictive allograft syndrome, mixed obstruction/restriction, and undefined, have been standardized, providing a framework for clinical practice and research (Table 1).[9] CLAD and PcGVHD share similar immune-mediated pathophysiologic mechanisms.[10] Therefore, we hypothesized that the ISHLT consensus definition for CLAD could be adapted to identify and characterize non-obstructive phenotypes of PcGVHD.

Methods

Patients

We retrospectively reviewed patient data from the "Natural History Study of Clinical and Biological Factors Determining Outcomes in Chronic Graft-Versus-Host Disease" (#NCT00092235), a study previously approved by the NIH Institutional Review Board and conducted in accordance with the Declaration of Helsinki. This study enrolled patients with cGVHD, referred by their primary HCT center to the NIH, where their pre- and post-HCT clinical data were collected comprehensively, and they underwent thorough clinical/laboratory workup and multidisciplinary evaluation.[11] The patients were followed for survival outcomes for as long as 10 years after the initial NIH visit. Consecutive patients enrolled between 1 October 2004 and 31 January 2020 were analyzed for study inclusion. Patients without a diagnosis of cGVHD or without complete PFT (spirometry, lung volume, and diffusion capacity) or CT imaging data were excluded. Follow-up data were obtained from the electronic medical records and censored on 1 June 2021.

Adapted ISHLT CLAD criteria and definition of study end points

We used an adapted criteria inspired by the 2019 ISHLT CLAD criteria[9] for PcGVHD adjudication (Table 1). Population reference values were used in the Adapted Criteria, instead of the subject's best PFTs. The adapted criteria defined PcGVHD as an FEV1 < 80% predicted in the absence of other etiologies. The criteria captured 4 PcGVHD subtypes based on their pulmonary function and CT scan characteristics: obstruction, restriction, mixed obstruction/restriction, and undefined (Table 1). Obstruction required obstructive ventilatory defect (FEV1/FVC < 0.7) without restriction (TLC > 90%) or restrictive CT findings. Restriction required restrictive ventilatory defect (TLC < 90% predicted) and restrictive CT findings. Mixed phenotype required mixed ventilatory defects and restrictive CT findings. The undefined was characterized by obstructive ventilatory defect plus either restrictive ventilatory defect or restrictive CT findings. Restrictive CT findings included any degree (patchy, unilobar, multilobar, or extensive) of ground glass opacities (GGO), parenchymal consolidation, traction bronchiectasis, lobar volume loss, usual interstitial pneumonia pattern, or pleural abnormalities.[12] Other abnormalities on CT scans, such as mosaic attenuation/air-trapping, cystic changes, or emphysema, were not restrictive-defining CT

Table 1. Comparison of the 2014 NIH cGVHD consensus criteria, the 2019 ISHLT CLAD criteria, and the adapted criteria

Criteria	NIH criteria	ISHLT CLAD criteria	Adapted criteria
Diagnosis	FEV1/VC < 0.7 or the 5th percentile predicted based on population-based reference; VC is either FVC or SVC, whichever is greater; FEV1 < 75% predicted with ≥ 10% decrease over less than 2 y, not corrected with albuterol	Persistent decline (> 3 mo, ≥ 20%) of FEV1 from the reference baseline; baseline is the mean of the best 2 post-transplant FEV1 measurements taken 3 wk apart	Abnormal pulmonary function after transplant (FEV1 < 80% predicted based on population-based reference), able to be classified into 1 of the 4 CLAD-PcGVHD subtypes, rule out other causes of pulmonary dysfunction
Phenotype	BOS: FEV1/VC < 0.7 or the 5th percentile predicted based on population-based reference; VC is either FVC or SVC, whichever is greater; evidence of air-trapping by expiratory CT or airway thickening or bronchiectasis by high-resolution CT, or air-trapping by PFT	BOS: obstruction (FEV1/FVC < 0.7), without restriction or CT opacity; RAS: restriction (TLC < 90% baseline) + CT opacity, FEV1/FVC ≥ 0.7; mixed: FEV1/FVC < 0.7, TLC < 90% baseline, with CT opacity; undefined: A. FEV1/FVC < 0.7, TLC < 90% baseline, NO CT opacity; B. FEV1/FVC < 0.7, TLC ≥ 90% baseline, WITH CT opacity	Obstruction: obstruction (FEV1/FVC < 0.7), without restrictive findings on PFT or CT; restriction: restriction (TLC < 90% predicted), with restrictive CT findings,* FEV1/FVC ≥ 0.7; mixed: FEV1/FVC < 0.7, TLC < 90% predicted, restrictive CT findings; undefined: A. FEV1/FVC < 0.7, TLC < 90% predicted, NO restrictive CT findings; B. FEV1/FVC < 0.7, TLC ≥ 90% predicted, WITH restrictive CT findings

RAS, restrictive allograft syndrome.

*Restrictive CT scan findings include ground glass opacities, parenchymal consolidation, traction bronchiectasis, lobar volume loss, usual interstitial pneumonitis pattern, and pleural abnormalities.

findings. For patients with multiple NIH-performed PFTs, the latest one was used for adjudication. Patients with normal post-HCT PFT(s) (FEV1 $\geq$ 80% predicted) were the control group in this study. Patients with abnormal post-HCT PFT(s) (FEV1 < 80% predicted) but lacked other findings to qualify for any PcGVHD phenotype, or with other pulmonary etiologies such as infection, were termed "unclassified" and did not meet the criteria for PcGVHD. The end points were post-HCT overall survival (OS), defined as the time from HCT to death from any cause, and post-cGVHD OS, defined as the time from the diagnosis of cGVHD to death from any cause.

Review and adjudication

An adjudication committee of 2 transplant pulmonologists, 1 transplant hematologist, and 1 cardiothoracic radiologist reviewed PFT, CT, and electronic medical records to adjudicate patients for PcGVHD using both the NIH[4] and the adapted criteria. We grouped the patients who met the adapted criteria into those who met the NIH criteria (NIH+) and those who did not (NIH−).

All NIH-performed PFTs were queried. For patients with FEV1 < 80% predicted, concomitant (within 1 month) thoracic CT images were separately reviewed by an independent cardiothoracic radiologist (A.S.). Two sets of independent reviewers (Y.P. and A.V.C. or Y.P. and M.B.K.) adjudicated subjects by the adapted criteria. Cases with inconsistent conclusion between the reviewers were discussed in a committee meeting (Y.P., A.V.C., M.B.K., and S.A.-E.) for consensus.

When available, pre-HCT and/or serial post-HCT PFT data were collected and compared with the adjudication PFT. We did not use the changes in the percentages of predictive values because of the anticipated change in the reference value over time with age, and different centers may use different reference values. The following equations were used to calculate the change in FEV1:

$$\Delta FEV1 \ (L) = \text{Pre-HCT or early post-HCT FEV1 (L)} \\ - \text{Adjudication FEV1 (L)}$$
$$\Delta FEV1, \% = \Delta FEV1 \ (L)/\text{Pre-HCT or early post-HCT} \\ FEV1 \ (L)$$

Statistical analysis

Demographic characteristics were described as counts and percentages for categorical variables and as mean and standard deviation (SD) for continuous variables. We used t test or Mann-Whitney U test, and analysis of variance or Kruskal-Wallis test to compare continuous variables with or without a normal distribution between 2 or more groups, respectively; χ^2 test was used for categorical variables. Pearson correlation analysis was used to analyze the correlation between FVC and TLC. Univariable and multivariable Cox proportional hazard models were used to evaluate the association of the clinical variables with post-HCT OS and post-cGVHD OS. The following variables were selected based on their plausible effect on HCT outcome: age, recipient sex, Karnofsky performance status (KPS), transplant location (at the NIH or in the community), transplant indication, donor type, graft source, and relapse/refractory malignancy.[11,13,14] Kaplan-Meier survival curves were shown for illustration purposes. All analysis were constructed using SAS 9.4 (SAS institute, Cary, NC). Statistical significance was defined as a 2-tailed $P \leq .05$.

Results

Study cohort

Of the 447 patients enrolled in the NIH cGVHD natural history study between 1 October 2004 and 31 January 2020, 97 were excluded (21 did not have cGVHD; 76 did not have an available PFT and/or CT), leaving 350 patients for the study cohort. The study cohort received HCT between 7 January 1987 and 21 March 2018. The median time from HCT to the diagnosis of cGVHD was 234 days (interquartile range [IQR], 144-373), and median time from cGVHD diagnosis to NIH enrollment was 795 days (IQR, 373-1491). A total of 117 patients had multiple NIH-performed PFTs (median number of PFT = 5; range, 2-35). The median time from HCT to the adjudication PFT was 1460 days (IQR, 948-2594 days).

In the study cohort, 260 (75.5%) had severe cGVHD (cGVHD severity score = 3). Among them, 132 (37.7%) were controls with normal post-HCT PFT. PcGVHD was diagnosed in 166 (47.4%) patients, including 12 (3.4%) obstruction, 67 (19.1%) restriction, 47 (13.4%) mixed, and 40 (11.4%) undefined. Fifty-two patients (14.9%) were in unclassified, including 34 (65.4%) with restrictive PFT but lacked CT abnormalities, 17 of which had chest skin sclerosis that likely explained the restrictive PFT. Chest wall sclerosis was also seen in 8 of the undefined, 9 of the mixed, and 31 of the restrictive subtype. Thirty-one of the restrictive and 9 of the mixed subtype showed PFT signs of intrapulmonary restriction with TLC < 90% and showed pulmonary imaging findings to support the restrictive PFT pattern. The 17 patients in the unclassified group with chest wall sclerosis all had reduced TLC but their CT chest showed no airway, parenchymal, or pleural abnormalities. Furthermore, diffusing capacity of carbon monoxide adjusted for hemoglobin (DLCO adj) was significantly lower in the patients with PcGVHD with chest wall sclerosis than the unclassified patients with chest wall sclerosis (51.9 $\pm$ 14.7% vs 61.5 $\pm$ 17.2, $P = .031$), indicative of a potentially greater degree of parenchymal involvement in the former. Another 9 (17.3%) of the unclassified group had decreased FEV1 but normal FEV1/FVC and normal TLC. Additionally, other etiologies for lung abnormalities, such as infection or primary disease, were identified in 15 unclassified patients.

Only 78 (47.0%) of the patients with adapted criteria-defined PcGVHD met the NIH criteria for BOS (NIH+). Reasons for not meeting the NIH criteria (88 patients, 53.0%, NIH−) were restrictive lung disease without obstruction (67, 76.1%), bronchodilator reversibility (11, 12.5%), or FEV1 $\geq$ 75% predicted (10, 11.4%). None of the patients in the control or unclassified group met the NIH criteria. Initial inter-rater agreement rate was 81.4% for whether PFT showed obstruction, restriction, or normal; inter-rater agreement rate for PcGVHD diagnosis and phenotype classification was 82.3%. The reason for the initial interrater disagreement was unfamiliarity with the criteria, and all the disagreements resolved after the adjudication meeting. The CONSORT flow diagram is shown in Figure 1.

Clinical characteristics of PcGVHD defined by the adapted criteria

The demographic features, HCT, and GVHD information of the 350 patients are shown in supplemental Table 1. Average age at HCT was 39.6 years (standard deviation [SD], 16.5). The majority (88.2%) were white. More than 75% of the patients had an NIH global score severe cGVHD, and most received more than 4 lines
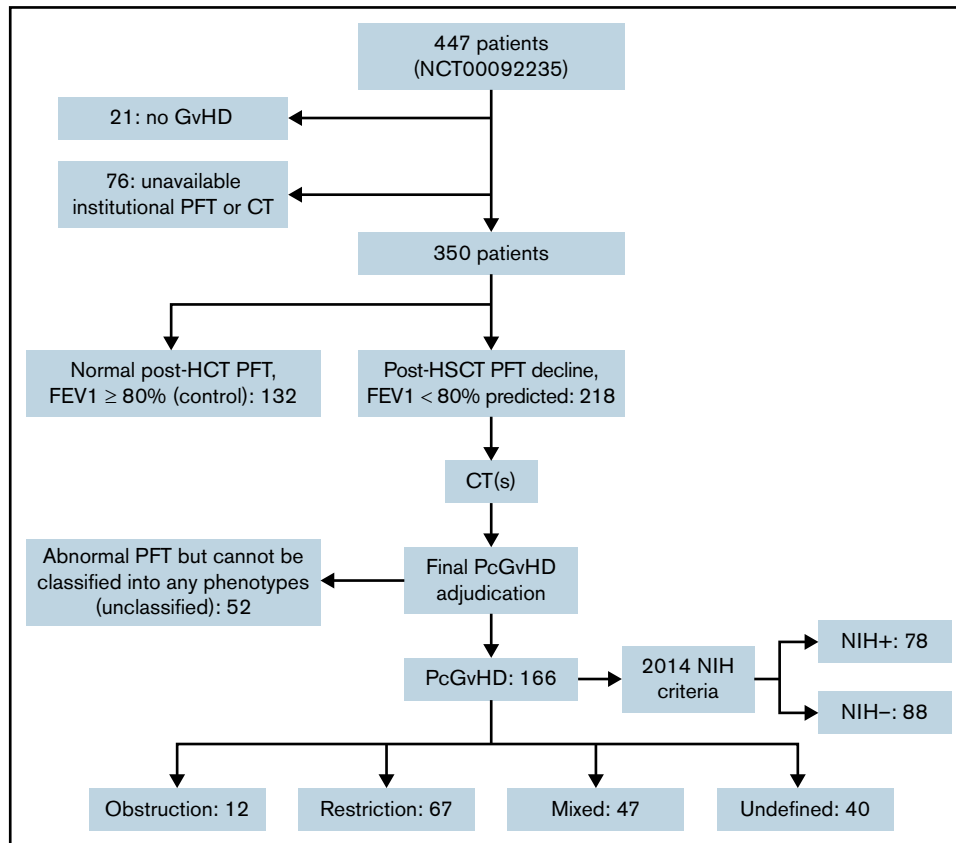
Figure 1. Consort flow diagram of the study cohort.

of systemic immunosuppressive therapy for cGVHD. Average KPS at the time of NIH visit was 78.3 (SD, 11.8).

Patients with PcGVHD had the lowest KPS compared with the control and unclassified groups (74.6 ± 12.0 vs 83.0 ± 9.7 and 77.6 ± 12.0, $P <$ .001), and as expected, the lowest FEV1 (50.3 ± 16.3 vs 97.4 ± 12.8 and 64.0 ± 14.2, $P <$ .001). More than 90% of patients in the PcGVHD group had severe cGVHD, whereas around 75% and less than 60% of patients in the unclassified and control groups, respectively, had severe cGVHD ($P <$ .001). The comparison of clinical characteristics among the control, PcGVHD, and unclassified groups is shown in Table 2.

Adapted criteria vs NIH criteria

Within the patients with adapted criteria–defined PcGVHD, the NIH− and NIH+ groups were similar with regard to the majority of clinical features. The NIH− group had a higher percentage of male patients and pre-HCT tobacco use than the NIH+ group. The NIH− group, mainly consisted of the restrictive phenotype, had significantly lower TLC and RV, and significantly higher FEV1, FEV1/FVC, and FEF 25% to 75% than the NIH+ group. DLCO was similar between the 2 groups. Although cGVHD severity were similar between the groups, skin, joints, and fascia cGVHD were more commonly seen in the NIH− group than the NIH+ group (Table 3).

Clinical features of different PcGVHD phenotypes

Within the 4 PcGVHD phenotypes, recipient sex, race, indication for HCT, conditioning intensity, rates of acute GVHD, and KPS

were similar (supplemental Table 2). Median time from HCT to diagnosis of cGVHD was more than 100 days later in patients who developed restrictive PcGVHD compared with those who develop obstructive or mixed PcGVHD ($P =$ .040). The time from HCT to adjudicative PFT were similar between the groups. Patients with obstruction were older at the time of transplant (average age, 51.8 years; SD, 13.1), and patients with undefined phenotype the youngest (average age, 30.8 years; SD, 17.4). Patients with the mixed phenotype more commonly had tobacco exposure before HCT. Busulfan exposure was more common in obstruction compared with others. Skin, joints, and fascia cGVHD coexisted with restriction in more than 80% of cases, significantly more than the other phenotypes. The rates of liver and gastrointestinal tract cGVHD were similar among the 4 phenotypes.

Although all patients with PcGVHD had pre-bronchodilator FEV1 < 80%predicted, the mixed phenotype had the lowest FEV1 (mean, 43.1%; SD, 15.0%); patients with obstruction and restriction had similar FEV1 (57.7 ± 15.9% vs 55.0 ± 14.1%, $P =$ .554). DLCO adj was the poorest in restriction and the best in obstruction (50.2 ± 17.8 vs 67.5 ± 21.0%, $P =$ .003) and not significantly different between the restriction and the mixed phenotypes ($P =$ .424). In the entire study cohort, TLC and FVC had a strong correlation ($R^2 =$ 0.809, $P <$ .001).

A myriad of thoracic CT findings was identified (Table 4; supplemental Figure 1); however, honeycombing or usual interstitial pneumonia pattern were absent. It was not uncommon for patients with restriction to have obstructive features on CT, such as mosaic

Table 2. Clinical features of the study cohort

Characteristics		Control (n = 132)	PcGVHD(n = 166)	Unclassified (n = 52)	P
Age at HCT*		39.8 ± 16.8	40.7 ± 15.5	35.3 ± 18.7	.114
Weight, kg*		72.0 ± 20.5	68.5 ± 16.7	67.1 ± 20.4	.170
KPS*		83.0 ± 9.7	74.6 ± 12.0	77.6 ± 12.0	<.001
TLC, %predicted*		95.3 ± 13.4	74.2 ± 17.9	76.4 ± 15.4	<.001
RV, %predicted*		88.1 ± 32.5	90.6 ± 42.0	89.8 ± 37.7	.845
FEV1/FVC, %, pre-bronchodilator*		79.4 ± 7.9	64.9 ± 18.1	76.3 ± 13.1	<.001
FVC, %predicted, pre-bronchodilator*		96.6 ± 12.4	62.7 ± 17.2	67.5 ± 15.0	<.001
FEV1, %predicted, pre-bronchodilator*		97.4 ± 12.8	50.3 ± 16.3	64.0 ± 14.2	<.001
FEF 25-75%, %predicted*		94.0 ± 27.9	39.9 ± 31.9	57.8 ± 26.5	<.001
DLCO adj, %predicted*		74.8 ± 18.4	53.7 ± 16.7	57.9 ± 16.7	<.001
Different types of cGVHD treatment*,†		3.8 ± 1.9	5.1 ± 2.2	5.2 ± 2.2	<.001
Time from HCT to cGVHD diagnosis (days)*		395.1 ± 741.0	319.7 ± 345.8	282.0 ± 208.7	.308
Time from HCT to adjudicative PFT (days)*		1822.8 ± 1640.6	2113.6 ± 1579.2	1869.6 ± 1345.4	.253
Recipient sex‡	Female	62 (46.9)	74 (44.6)	20 (38.4)	.579
Recipient race‡	White	112 (86.2)	148 (89.2)	47 (90.4)	.635
HCT location‡	Local	113 (85.6)	154 (92.8)	45 (86.5)	.115
	NIH	19 (14.4)	12 (7.2)	7 (13.5)	
Pre-HCT tobacco use‡		34 (25.9)	51 (30.7)	12 (23.1)	.471
HCT indication‡	AML, MDS, ALL	61 (46.2)	96 (57.8)	26 (50.0)	.037
	CML, MPN	16 (12.1)	19 (11.5)	3 (5.8)	
	CLL, HL, NHL, MM	48 (36.4)	43 (25.9)	15 (28.9)	
	Others	7 (5.3)	8 (4.8)	8 (15.4)	
Disease status at HCT‡	CR	47 (39.8)	92 (60.1)	21 (44.7)	.003
HLA‡	MRD	69 (52.7)	90 (54.2)	18 (36.7)	.037
	MUD	40 (30.5)	53 (31.9)	28 (57.1)	
	MMUD	16 (12.2)	16 (9.6)	2 (4.1)	
	Haploidentical	6 (4.6)	7 (4.2)	1 (2.0)	
Stem cell source‡	PB	98 (77.2)	138 (84.2)	38 (73.1)	.233
	BM	29 (22.8)	26 (15.9)	12 (23.1)	
Conditioning‡	RIC	71 (53.8)	69 (42.1)	25 (50.0)	.126
	MAC	61 (46.2)	95 (57.9)	25 (50.0)	
Busulfan‡		44 (34.4)	76 (48.1)	16 (32.0)	.026
TBI‡		54 (41.2)	64 (39.0)	20 (39.2)	.924
GVHD prophylaxis	CNI + MTX	62 (50.8)	70 (46.1)	29 (59.2)	.269
Second malignancy‡		20 (15.1)	34 (20.5)	10 (19.2)	.488
Acute GVHD‡		92 (69.7)	120 (72.3)	38 (73.1)	.851
cGVHD global	Mild	6 (4.5)	1 (0.6)	0	<.001
severity‡	Moderate	50 (37.9)	15 (9.0)	13 (25.0)	
	Severe	76 (57.6)	150 (90.4)	39 (75.0)	
Relapse/refractory primary malignancy‡		22 (16.7)	22 (13.3)	8 (15.4)	.708

ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BM, bone marrow; CLL, chronic lymphocytic leukemia; CML, chronic myelogenous leukemia; CR, complete remission; HL, Hodgkin's lymphoma; MAC, myeloablative conditioning; MDS, myelodysplastic syndrome; MM, multiple myeloma; MMUD, mismatched unrelated donor; MPN, myeloproliferative neoplasm; MRD, matched related donor; MUD, matched unrelated donor; NHL, non-Hodgkin's lymphoma; PB, peripheral blood; RIC, reduced intensity conditioning; RV, residual volume; TBI, total body irradiation.

*Mean ± SD, analysis of variance test.
†Excluding topical treatment such as topical corticosteroids.
‡Count (percentage), χ^2 test.

Table 3. Comparison of the clinical features between the NIH− and the NIH+ groups

Characteristics		NIH− (n = 88)	NIH+ (n = 78)	P
Age at HCT*		39.6 ± 15.5	42.0 ± 15.45	.325
Weight, kg*		70.9 ± 18.0	65.8 ± 14.88	.050
KPS*		75.1 ± 12.3	74.0 ± 11.6	.551
TLC, %predicted*		69.8 ± 17.0	79.1 ± 17.7	<.001
RV, %predicted*		83.8 ± 34.2	98.3 ± 48.4	.027
FVC, %predicted, pre-bronchodilator†		59.5 ± 17.8	66.4 ± 15.8	.009
FEV1, %predicted, pre-bronchodilator*		56.6 ± 14.9	43.3 ± 15.0	<.001
FEF 25-75%, %predicted†		58.4 ± 33.2	19.1 ± 10.0	<.001
FEV1/FVC, %, pre-bronchodilator†		76.7 ± 12.3	51.6 ± 13.8	<.001
DLCO adj, %predicted*		53.5 ± 17.9	54.0 ± 15.3	.848
Time from HCT to cGVHD diagnosis (days)*		361.0 ± 332.2	273.2 ± 357.0	.103
Time from HCT to adjudicative PFT (days)*		2178.0 ± 1724.1	2040.9 ± 1405.4	.578
Different types of cGVHD treatment*,‡		3.8 ± 1.9	5.1 ± 2.2	<.001
Recipient sex§	Male	57 (64.7)	35 (44.8)	.010
Recipient race§	White	78 (88.6)	70 (89.7)	.819
HCT location§	Local	81 (92.0)	73 (93.6)	.701
	NIH	7 (8.0)	5 (6.4)	
Pre-HCT tobacco use§		68 (77.3)	47 (60.3)	.018
HCT indication§	AML, MDS, ALL	47 (53.4)	49 (62.8)	.421
	CML, MPN	9 (10.2)	10 (12.8)	
	CLL, HL, NHL, MM	27 (30.7)	16 (10.5)	
	Others	5 (5.7)	3 (3.8)	
Disease status at HCT§	CR	46 (52.3)	46 (59.0)	.619
HLA§	MRD	48 (54.5)	42 (53.8)	.937
	MUD	29 (33.0)	24 (30.8)	
	MMUD	8 (9.1)	8 (10.3)	
	Haploidentical	3 (3.4)	4 (5.1)	
Stem cell source§	PB	69 (78.4)	69 (88.5)	.150
	BM	17 (19.3)	9 (11.5)	
Conditioning§	RIC	39 (44.3)	30 (38.5)	.448
	MA	48 (54.5)	47 (60.3)	
Busulfan§		40 (45.5)	36 (46.2)	.662
TBI§		36 (40.9)	28 (38.9)	.924
GVHD prophylaxis	CNI + MTX	44 (50.0)	38 (48.7)	.921
Second malignancy§		18 (20.5)	16 (20.5)	.993
Acute GVHD§		23 (26.1)	23 (29.5)	.630
cGVHD global severity§	Mild	1 (1.1)	0	.336
	Moderate	10 (11.4)	5 (6.4)	
	Severe	77 (87.5)	73 (93.6)	
Skin cGVHD§		74 (84.1)	54 (69.2)	.005
Joints/fascia cGVHD§		68 (77.3)	45 (57.7)	.007
Liver cGVHD§		49 (55.7)	40 (51.3)	.570
GI tract cGVHD§		39 (44.3)	34 (43.6)	.925
Relapse/refractory primary malignancy§		11 (12.5)	11 (14.1)	.761

ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BM, bone marrow; CLL, chronic lymphocytic leukemia; CML, chronic myelogenous leukemia; CR, complete remission; GC, GVHD control; HL, Hodgkin's lymphoma; MAC, myeloablative conditioning; MDS, myelodysplastic syndrome; MM, multiple myeloma; MMUD, mismatched unrelated donor; MPN, myeloproliferative neoplasm; MRD, matched related donor; MUD, matched unrelated donor; NHL, non-Hodgkin's lymphoma; PB, peripheral blood; RIC, reduced intensity conditioning; RV, residual volume; TBI, total body irradiation.

*Mean ± SD, independent sample t test.
†Mean ± SD, Mann-Whitney U test.
‡Excluding topical treatment such as topical corticosteroids.
§Count (percentage), χ^2 test.

Table 4. Thoracic CT findings in different PcGVHD phenotypes

Thoracic CT scan findings		Obstruction (N = 12)	Restriction (N = 67)	Mixed (N = 47)	Undefined (N = 40)	P
Obstructive features*	Mosaic attenuation	6 (50.0)	34 (50.7)	39 (82.9)	29 (72.5)	.002
	Air-trapping	6 (50.0)	36 (53.7)	37 (78.7)	25 (62.5)	.040
	Airway thickening	4 (33.3)	17 (25.4)	29 (61.7)	16 (40.0)	.001
	Non-traction bronchiectasis	2 (16.7)	20 (29.9)	24 (51.1)	26 (65.0)	<.001
	Cystic changes	0	5 (7.5)	2 (4.3)	2 (5.0)	.712
	emphysema	0	2 (3.0)	3 (6.4)	0	.328
Restrictive	GGO	0	27 (40.3)	31 (66.0)	7 (17.5)	<.001
features*,†	Consolidation	0	8 (11.9)	9 (19.1)	1 (2.5)	.055
	Traction bronchiectasis	0	23 (34.3)	11 (23.4)	3 (7.5)	.007
	Lobar volume loss	0	20 (29.9)	9 (19.1)	1 (2.5)	.003
	Pleural abnormalities	0	29 (43.3)	16 (34.0)	4 (10.0)	.002
	Reticulation	0	28 (41.8)	19 (40.4)	8 (40.0)	.050
	PPFE	0	3 (4.5)	1 (2.1)	1 (2.5)	.748
Normal*		6 (50.0%)	0	0	6 (15.0)	<.001

PPFE, pleuroparenchymal fibroelastosis.

*Number (percentage), χ^2 test.

†χ^2 test between the RLS, M, and UD groups, without the BOS group.

attenuation pattern (50.7%) and air-trapping (53.7%). Traction bronchiectasis, lobar volume loss, and pleural abnormalities were more commonly seen in the isolated restrictive phenotype. The mixed phenotype had high rates of both obstructive and restrictive CT features, the most common being mosaic attenuation pattern (82.9%), air-trapping (78.7%), and GGO (66.0%).

Association between PcGVHD and OS

The median post-HCT OS of the study cohort was 203 months (range, 8-392 months). Patients with PcGVHD had an increased risk of death compared with those without PcGVHD (multivariable Cox model, see Methods; hazard ratio [HR] = 1.74, 95% confidence interval [CI] = 1.11-2.72, P = .015). OS of the cGVHD control and unclassified groups was similar (P = .668; Figure 2A). The median post-HCT OS of the NIH− and NIH+ PcGVHD groups was similar, 172 months (IQR = 81-230) vs 165 months (IQR = 84-214; HR = 0.89, 95% CI = 0.53-1.52, P = .678), both significantly poorer than patients without PcGVHD (HR = 1.75, 95% CI = 1.07-2.84, P = .025 and HR = 1.88, 95% CI = 1.20-2.95, P = .006, respectively; Figure 2B). Among different PcGVHD phenotypes, the restriction and undefined groups had worse survival than the control group (HR = 1.93, 95% CI = 1.14-3.27, P = .015, and HR = 2.25, 95% CI = 1.24-4.14, P = .009, respectively), whereas survival between the obstruction or mixed group vs the control group was not significantly different (HR = 0.38, 95% CI = 0.05-2.86, P = .346 and HR = 1.49, 95% CI = 0.83-2.65, P = .181, respectively; Figure 2C).

After cGVHD diagnosis, the median OS of the study cohort was 184 months (range, 3-388 months). Similar to post-HCT OS, patients with PcGVHD had a higher risk of death compared with controls (HR = 1.60, 95% CI = 1.03-2.48, P = .037), whereas patients in the control and unclassified groups had similar risk of death (P = .413; supplemental Figure 2a). NIH− and NIH+ PcGVHD groups showed similar risks of death after cGVHD diagnosis (HR = 0.86, 95% CI = 0.51-1.46, P = .578; supplemental

Figure 2b). Similarly, the restriction and undefined groups had the worst post-cGVHD diagnosis survival among different PcGVHD phenotypes (supplemental Figure 2c).

Longitudinal trend of pulmonary function after HCT

Pretransplant PFT results were available in 84 patients (24%) and early (within 1 year of HCT) post-HCT PFTs were available in 21 (6%) of the study cohort. Only 54 had volume data, making the calculation of ΔFEV1 possible; the rest only had percentage of the populational reference value. Various degrees of declination in FEV1 were seen in the patients with PcGVHD, with mean ΔFEV1 −1.37 to −1.13 L and ΔFEV1% −51 to −35% (supplemental Table 3). Conversely, the declination in the unclassified and control groups were less prominent, with mean ΔFEV1 −0.42 to −0.07 L and ΔFEV1% −14 to +1%.

One patient in the restrictive group had abnormal PFT pre-HCT with further decline after HCT. Her chest CT scan demonstrated a pattern of pleuroparenchymal elastosis (PPFE), a rare entity characterized by upper lobe-predominant pleural and subpleural parenchymal fibrosis, seen in both lung transplant and HCT (Figure 3).[15,16]

Discussion

Informed by the potential similarities in immunopathogenesis for both PcGVHD after HCT and CLAD after lung transplantation, we refined and classified PcGVHD phenotypes based on the adapted criteria inspired by the ISHLT CLAD consensus criteria in a large cohort of patients with cGVHD. We intended to systematically evaluate post-HCT PFTs and thoracic CTs to novel PcGVHD phenotypes with clinical features that are distinct from patients with obstruction, some of which were not captured by the NIH cGVHD consensus criteria. Different from the NIH criteria that used post-bronchodilator FEV1 and FVC, the adapted criteria used pre-bronchodilator values, as airway smooth muscle tone could play a significant role in the pathophysiology in BOS, and bronchodilator
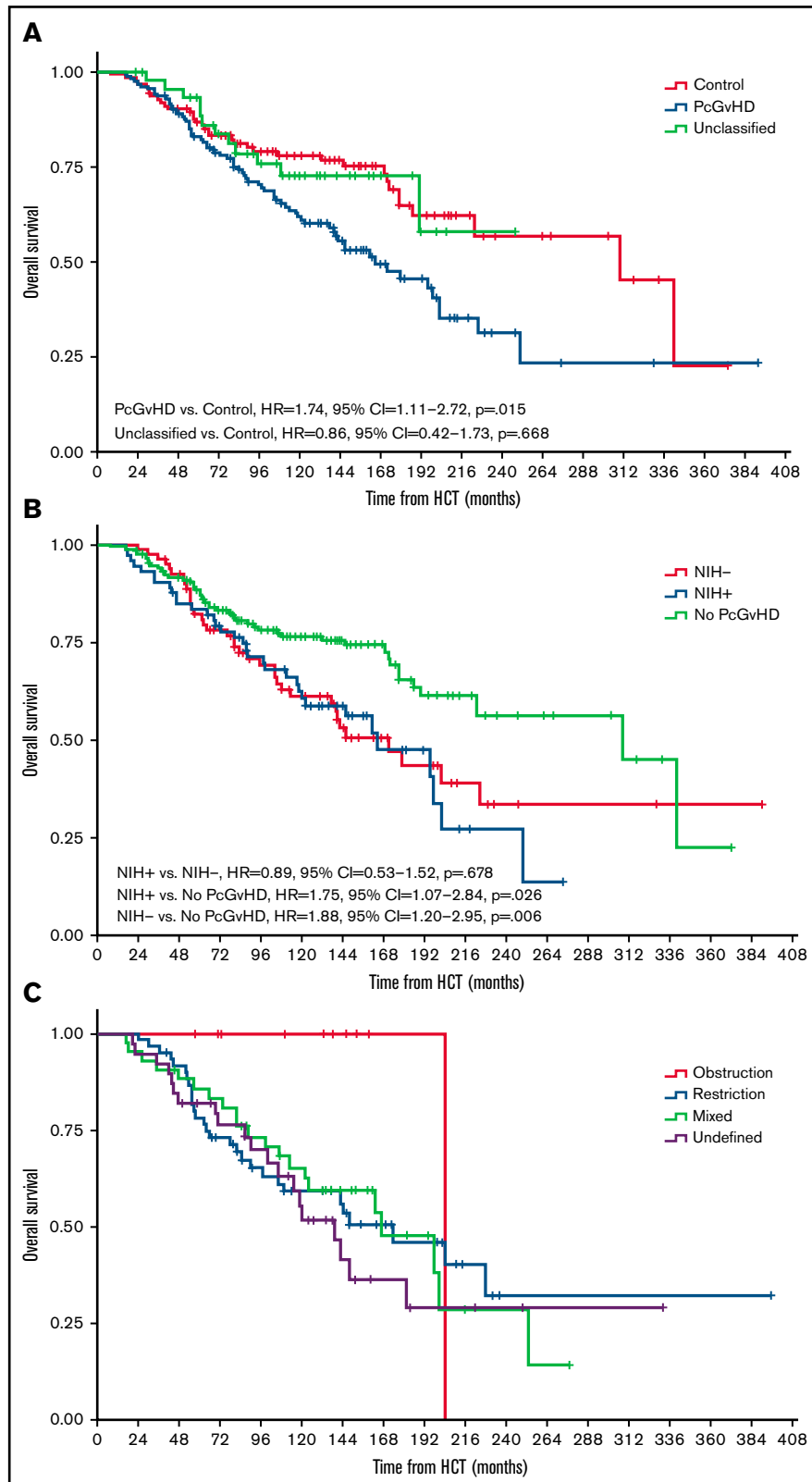
Figure 2. Kaplan-Meier curves displaying the post-HCT OS in the study cohort. (A) Comparison of post-HCT OS among the control, PcGVHD, and unclassified groups. (B) Comparison of post-HCT OS between the NIH− and NIH+ groups and those who did not meet the adapted criteria. (C) Comparison of post-HCT OS among different PcGVHD phenotypes. All hazard ratios (95% CI, *P* values) noted in the figures were from multivariable Cox proportional hazard models adjusting for age, recipient sex, KPS, transplant location (at the NIH or in the community), transplant indication, donor type, graft source, and relapse/refractory malignancy.
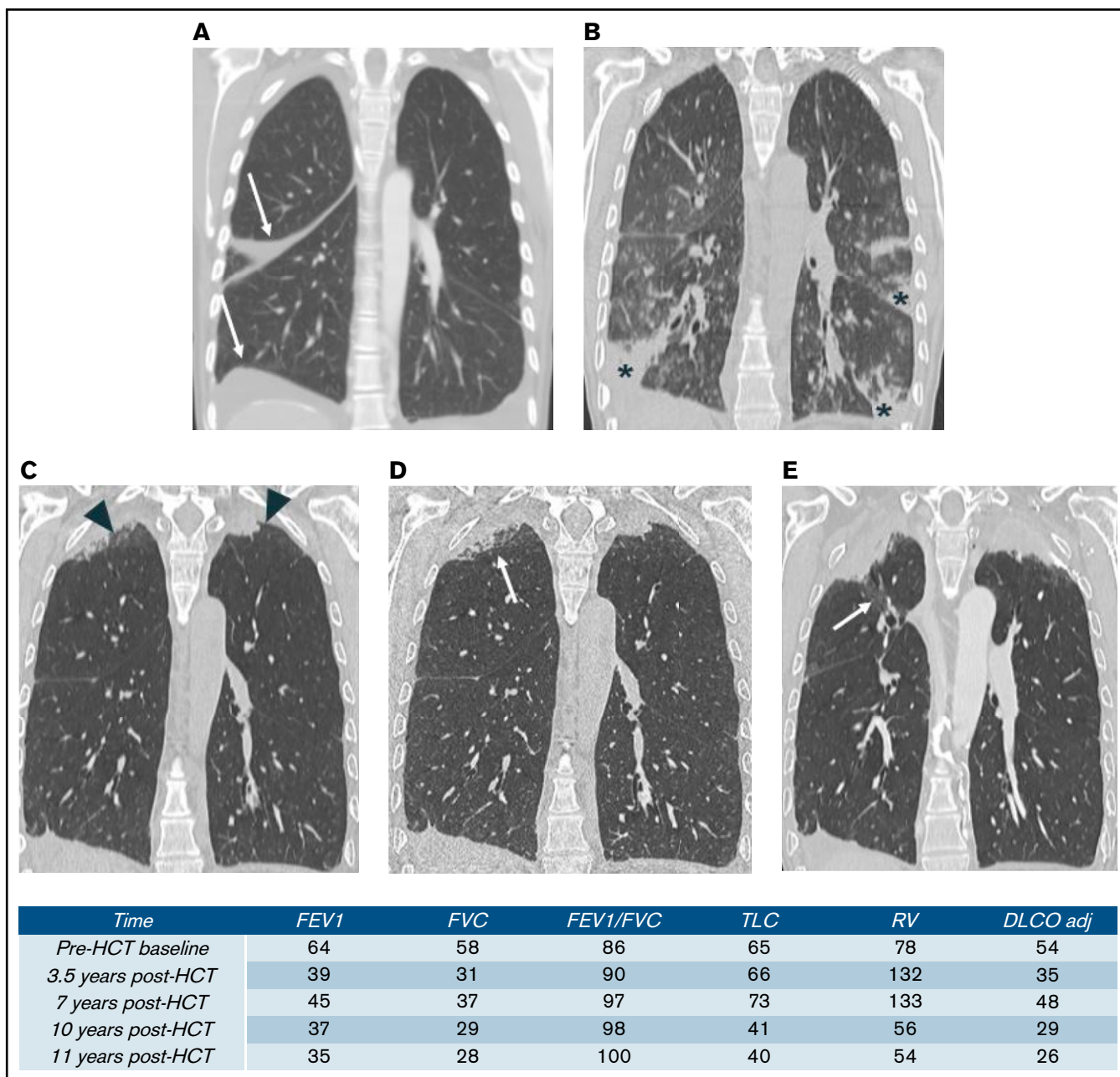
Figure 3. Serial thoracic CT of a patient with restrictive PcGVHD. (A) Baseline, with small pleural effusion (arrow). (B) At 3.5 years after HCT, admitted for hypoxia. CT scan shows normal lung apices and patchy foci of consolidation within the mid- to lower lungs (stars). (C-D) At 7 years after HCT. (C) Routine CT scan and (D) high-resolution CT scan images show apical pleural thickening (arrowhead) and subpleural consolidation and reticulation (arrow) that are associated with upper lobe traction bronchiectasis and volume loss. (E) At 10 years after HCT. CT scan shows progressive upper lobe volume loss, pleural thickening, and subpleural fibrotic consolidation. Upper lobe traction bronchiectasis (arrow) has also progressed.

reversibility could not exclude BOS.[17] The TLC cutoff of 90% was used, as a previous study has shown that the combination of TLC < 90% and interstitial changes on the thoracic CT was sensitive and specific for diagnosing interstitial disease after HCT, with positive and negative predictive values 1.00 and 0.75, respectively.[18] Using this adapted criteria, we reported that about half of patients with severe cGVHD potentially have PcGVHD; less than half of these patients with PcGVHD met the NIH criteria. Much like patients with PcGVHD who met NIH criteria (NIH+), the patients

with PcGVHD uncaptured by the NIH criteria (NIH−) showed worse survival compared with cGVHD controls without PcGVHD. The adapted criteria may therefore provide more informative clinical phenotypes and risk stratification of PcGVHD.

Multiple prior case reports and case series have reported nonobstructive or coexisting obstructive and restrictive late pulmonary complications in HCT recipients.[16] In a single center study, 31 of 1277 HCT recipients (2.4%) were diagnosed with restriction and

interstitial lung disease, with median time from HCT to diagnosis of 11.3 months.[19] In a case series of 70 patients who received lung transplantation post-HCT between 1990 and 2013, 52 (74.3%) of the explanted lung had BOS, and 17 (24.3%) had pulmonary interstitial fibrosis.[20] In another study involving 60 patients that received lung transplantation after HCT, 36 met the 2014 NIH criteria for BOS and 24 (40.0%) did not. On histology, BOS occurred in 55 (92%) and pulmonary fibrosis occurred in 34 (57%) of patients, and the 2 coexisted in 29 (48%) of patients.[18] Another study focusing on late onset post-HCT noninfectious pulmonary complications showed that, at a median follow up of 72.3 (15.2-88.5) months, 22 (11.1%) and 12 (6.0%) of 198 HCT recipients were diagnosed with BOS and interstitial lung disease, respectively.[21] In our cohort, isolated restrictive lung disease was the most common phenotype detected compared with obstruction and mixed obstruction/restriction, a higher prevalence than others have previously reported in literature, which could be related to the difference between our cohort, which were patients referred from various HCT centers around the United States vs other single center series.[8,18-21] Despite the differences, our results, together with others, highlight the limitation of the current NIH criteria. Given the high prevalence of pulmonary dysfunction in patients with severe cGVHD, it is worthwhile to develop a more diligent, lung transplant–like monitoring strategy of pulmonary function in these patients, and methods such as home spirometry monitoring are worth exploring.[22]

We confirmed in our cohort that similar phenotypes of pulmonary disease are observed in cGVHD after HCT and CLAD after lung transplantation. We demonstrated the strong correlation between FVC and TLC, which was also seen in CLAD, supporting the use FVC as an alternative method to monitor post-HCT lung volume in resource-limiting situations.[23] Discrepancies remained in the distribution of different phenotypes between our study and the CLAD literature. A single center study of 506 lung transplant recipients showed that 174 patients developed CLAD, including 104 (59.8%) with obstruction, whereas restriction, mixed, and undefined phenotypes each comprised around 10% of patients.[23] In contrast, our cohort after HCT had a high percentage of restriction, mixed, and undefined, whereas obstruction comprised only 7.2% of the PcGVHD population. The differences between our results and lung transplantation literature may again reflect the difference between our multicenter-based referral cohort vs single center cohorts but also may reflect differences in pathophysiology between PcGVHD and CLAD. In cGVHD, there may be more wide-spread immune-mediated tissue injury and remodeling, whereas in lung transplantation, there could be ischemia-reperfusion injury early after surgery that contributes to CLAD.[24] Moreover, in CLAD, the patient's best posttransplant PFT is used as the baseline to monitor PFT decline, whereas the current study used the populational predictive value.

The small size of the obstructive cohort may not accurately reflect the outcome of this phenotype, which may explain the discrepancy between our survival results with the existing CLAD and GVHD BOS literature.[5,6] Other reasons for the relatively better outcome of the obstructive phenotype observed in our cohort, in comparison with post-HCT BOS literature, could be that the prior BOS literature included all patients with obstructive components (ie, the combination of obstructive, mixed, and undefined), whereas in the current study these were analyzed separately. Future prospective studies

with a larger cohort of obstructive PcGVHD is needed to determine the relative survival of different PcGVHD phenotypes, in comparison with the CLAD counterparts in lung transplantation.

The major limitation of our study is the lack of baseline and chronological PFT and CT data in close to 80% of the study cohort. Around 16% to 17% of patients may have pre-HCT PFT abnormality, with FEV1 or DLCO < 80% predicted; pre-HCT restrictive lung disease has been reported in as high as 7.6% of patients.[25,26] The rate of PcGVHD could have been overestimated in our study because of the lack of comparison with pre-HCT PFT. Because of this limitation, we could not establish the accurate time of PcGVHD onset and could not analyze the survival after PcGVHD diagnosis. Furthermore, we could not implement a "decline from baseline" criterion similar to the NIH criteria or the ISHLT-CLAD criteria, in the current adapted criteria. Nonetheless, in patients with available PFT volume data, the patients who met the current adapted criteria all had chronological decline. In future studies using cohorts with longitudinal PFT monitoring, we would like to test a modified adapted criteria with the addition of a "decline from baseline" requirement, which will likely improve the performance of the criteria in diagnostic accuracy.

Other limitations should be noted in our study. First, the number of haploidentical HCT was very small and may not reflect the current landscape of graft source and posttransplant cyclophosphamide-based GVHD prophylaxis. Second, patients were referred and most had severe GVHD, rendering the cohort liable to referral and selection bias, and therefore, potentially not representative of general cGVHD population. Third, the cohort lacked non-GVHD HCT recipients and could not compare the characteristics of patients with chronic restrictive lung disease without cGVHD vs patient with PcGVHD defined by the adapted criteria, nor could it establish the association of restrictive lung disease with cGVHD. Restrictive lung disease in the absence of GVHD has been described in HCT recipients.[19] Fourth, lung volume was measured with a dilutional method (nitrogen washout) in our study, which may underestimate the TLC and have inadvertently classified some patients with purely obstructive phenotype as mixed phenotype.[27] Last, there was insufficient data regarding donor-recipient's blood group matching and cytomegalovirus status, and it would be worthwhile to evaluate the effects of these factors on the development of PcGVHD in future studies.

In conclusion, with PcGVHD criteria adapted from the 2019 ISHLT CLAD Consensus guidelines, we identified novel phenotypes of pulmonary dysfunction after HCT with varying implications on survival, supporting further investigation into their relationship with cGVHD and a re-classification of the current NIH criteria. The adapted criteria demonstrated potential to effectively and systematically identify different phenotypes of chronic lung disease in patients with cGVHD. The current limitations of the adapted criteria, namely the failure to take into consideration the trend of pulmonary function decline after HCT, will need to be overcome in future longitudinal studies. The similarities between PcGVHD and CLAD can facilitate more cross-talks between lung transplantation and HCT field.

Acknowledgments

References

1. Zeiser R, Blazar BR. Pathophysiology of chronic graft-versus-host disease and therapeutic targets. *N Engl J Med.* 2017;377(26):2565-2579.

2. Csanadi M, Agh T, Tordai A, et al. A systematic literature review of incidence, mortality, and relapse of patients diagnosed with chronic graft versus host disease. *Expert Rev Hematol.* 2019;12(5):311-323.

3. Bergeron A. Late-onset noninfectious pulmonary complications after allogeneic hematopoietic stem cell transplantation. *Clin Chest Med.* 2017; 38(2):249-262.

4. Jagasia MH, Greinix HT, Arora M, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. The 2014 Diagnosis and Staging Working Group report. *Biol Blood Marrow Transplant.* 2015;21(3):389-401.e1.

5. Au BKC, Au MA, Chien JW. Bronchiolitis obliterans syndrome epidemiology after allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant.* 2011;17(7):1072-1078.

6. Ditschkowski M, Elmaagacli AH, Koldehoff M, Gromke T, Trenschel R, Beelen DW. Bronchiolitis obliterans after allogeneic hematopoietic SCT: further insight: new perspectives? *Bone Marrow Transplant.* 2013;48(9):1224-1229.

7. Hofstetter E, Henig N, Kappeler D. Prevalence of bronchiolitis obliterans syndrome (BOS) following allogeneic hematopoietic stem cell transplant (alloHSCT) in the USA, Europe and Japan. *Blood.* 2019;134(suppl 1):5678.

8. Namkoong H, Ishii M, Mori T, et al. Clinical and radiological characteristics of patients with late-onset severe restrictive lung defect after hematopoietic stem cell transplantation. *BMC Pulm Med.* 2017;17(1):123.

9. Verleden GM, Glanville AR, Lease ED, et al. Chronic lung allograft dysfunction: definition, diagnostic criteria, and approaches to treatment-A consensus report from the Pulmonary Council of the ISHLT. *J Heart Lung Transplant.* 2019;38(5):493-503.

10. Chien JW, Duncan S, Williams KM, Pavletic SZ. Bronchiolitis obliterans syndrome after allogeneic hematopoietic stem cell transplantation-an increasingly recognized manifestation of chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2010;16(1 suppl):S106-S114.

11. Kerep AZ, Broome J, Pirsl F, et al. Impact of the 2014 NIH chronic graft-versus-host disease scoring criteria modifications assessed in a large cohort of severely affected patients. *Bone Marrow Transplant.* 2019;54(1):76-84.

12. Glanville AR, Verleden GM, Todd JL, et al. Chronic lung allograft dysfunction: definition and update of restrictive allograft syndrome-A consensus report from the Pulmonary Council of the ISHLT. *J Heart Lung Transplant.* 2019;38(5):483-492.

13. Della Porta MG, Alessandrino EP, Bacigalupo A, et al; Gruppo Italiano Trapianto di Midollo Osseo. Predictive factors for the outcome of allogeneic transplantation in patients with MDS stratified according to the revised IPSS-R. *Blood.* 2014;123(15):2333-2342.

14. Finke J, Schmoor C, Bethge WA, et al; ATG-Fresenius Trial Group. Prognostic factors affecting outcome after allogeneic transplantation for hematological malignancies from unrelated donors: results from a randomized trial. *Biol Blood Marrow Transplant.* 2012;18(11):1716-1726.

15. Fujikura Y, Kanoh S, Kouzaki Y, Hara Y, Matsubara O, Kawana A. Pleuroparenchymal fibroelastosis as a series of airway complications associated with chronic graft-versus-host disease following allogeneic bone marrow transplantation. *Intern Med.* 2014;53(1):43-46.

16. Meignin V, Thivolet-Bejui F, Kambouchner M, et al. Lung histopathology of non-infectious pulmonary complications after allogeneic haematopoietic stem cell transplantation. *Histopathology.* 2018;73(5):832-842.

17. Barisione G, Bacigalupo A, Crimi E, Brusasco V. Acute bronchodilator responsiveness in bronchiolitis obliterans syndrome following hematopoietic stem cell transplantation. *Chest.* 2011;139(3):633-639.

18. Greer M, Riise GC, Hansson L, et al. Dichotomy in pulmonary graft-*versus*-host disease evident among allogeneic stem-cell transplant recipients undergoing lung transplantation. *Eur Respir J.* 2016;48(6):1807-1810.

19. Schlemmer F, Chevret S, Lorillon G, et al. Late-onset noninfectious interstitial lung disease after allogeneic hematopoietic stem cell transplantation. *Respir Med.* 2014;108(10):1525-1533.

20. Cheng G-S, Edelman JD, Madtes DK, Martin PJ, Flowers MED. Outcomes of lung transplantation after allogeneic hematopoietic stem cell transplantation. *Biol Blood Marrow Transplant.* 2014;20(8):1169-1175.

21. Bergeron A, Chevret S, Peffault de Latour R, et al. Noninfectious lung complications after allogeneic haematopoietic stem cell transplantation. *Eur Respir J.* 2018;51(5):1702617.

22. Morlion B, Knoop C, Paiva M, Estenne M. Internet-based home monitoring of pulmonary function after lung transplantation. *Am J Respir Crit Care Med.* 2002;165(5):694-697.

23. Levy L, Huszti E, Renaud-Picard B, et al. Risk assessment of chronic lung allograft dysfunction phenotypes: validation and proposed refinement of the 2019 International Society for Heart and Lung Transplantation classification system. *J Heart Lung Transplant.* 2020;39(8):761-770.

24. Watanabe T, Martinu T, Chruscinski A, et al. A B cell-dependent pathway drives chronic lung allograft rejection after ischemia-reperfusion injury in mice. *Am J Transplant.* 2019;19(12):3377-3389.

25. Ramirez-Sarmiento A, Orozco-Levi M, Walter EC, Au MA, Chien JW. Influence of pretransplantation restrictive lung disease on allogeneic hematopoietic cell transplantation outcomes. *Biol Blood Marrow Transplant.* 2010;16(2):199-206.

26. Parimon T, Madtes DK, Au DH, Clark JG, Chien JW. Pretransplant lung function, respiratory failure, and mortality after stem cell transplantation. *Am J Respir Crit Care Med.* 2005;172(3):384-390.

27. Tantucci C, Bottone D, Borghesi A, Guerini M, Quadri F, Pini L. Methods for measuring lung volumes: is there a better one? *Respiration.* 2016; 91(4):273-280.