ORIGINAL ARTICLE

AASLD
AMERICAN ASSOCIATION FOR
THE STUDY OF LIVER DISEASES

Abnormal liver tests are not sufficient for diagnosis of hepatic graft-versus-host disease in critically ill patients

Alexander H. Yang[1] [iD] | Mai Ai Thanda Han[1] | Niharika Samala[1] | Bisharah S. Rizvi[1] | Rachel Marchalik[2] | Ohad Etzion[1] | Elizabeth C. Wright[3] | Ruchi Patel[1] | Vinshi Khan[1] | Devika Kapuria[1] | Vikramaditya Samala Venkat[1] | David E. Kleiner[4] | Christopher Koh[1] | Jennifer A. Kanakry[2] | Christopher G. Kanakry[2] | Steven Pavletic[5] | Kirsten M. Williams[6] | Theo Heller[1]

[1]Liver Diseases Branch, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), Bethesda, Maryland, USA

[2]Experimental Transplantation and Immunotherapy Branch, National Cancer Institute (NCI), Bethesda, Maryland, USA

[3]Office of the Director, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), Bethesda, Maryland, USA

[4]Laboratory of Pathology, National Cancer Institute (NCI), Bethesda, Maryland, USA

[5]Immune Deficiency Cellular Therapy Program, National Cancer Institute (NCI), Bethesda, Maryland, USA

[6]Department of Pediatrics, Aflac Cancer and Blood Disorders Center, Children's Healthcare of Atlanta, Emory University, Atlanta, Georgia, USA

Correspondence
Alexander H. Yang and Theo Heller, Liver Disease Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 10, Rm 10N248A, 9000 Rockville Pike, Bethesda, MD 20892, USA.
Email: alexander.yang2@nih.gov and theoh@intra.niddk.nih.gov

Funding information
National Cancer Institute, National Institute of Diabetes and Digestive and Kidney Diseases Intramural Research Program, National Institutes of Health.

Abstract

Hepatic graft-versus-host disease (HGVHD) contributes significantly to morbidity and mortality after hematopoietic stem cell transplantation (HSCT). Clinical findings and liver biomarkers are neither sensitive nor specific. The relationship between clinical and histologic diagnoses of HGVHD was assessed premortem and at autopsy. Medical records from patients who underwent HSCT at the National Institutes of Health (NIH) Clinical Center between 2000 and 2012 and expired with autopsy were reviewed, and laboratory tests within 45 days of death were divided into 15-day periods. Clinical diagnosis of HGVHD was based on Keystone Criteria or NIH Consensus Criteria, histologic diagnosis based on bile duct injury without significant inflammation, and exclusion of other potential etiologies. We included 37 patients, 17 of whom had a cholestatic pattern of liver injury and two had a mixed pattern. Fifteen were clinically diagnosed with HGVHD, two showed HGVHD on autopsy, and 13 had histologic evidence of other processes but no HGVHD. Biopsy or clinical diagnosis of GVHD of other organs during life did not correlate with HGVHD on autopsy. The diagnostic accuracy of the current criteria was poor ($\kappa = -0.20$). A logistic regression model accounting for dynamic changes included peak bilirubin 15 days before death, and an increase from period −30 (days 30 to 16 before death) to period −15 (15 days before death) showed an area under the receiver operating characteristic curve of 0.77. Infection was the immediate cause of death in 68% of patients. In conclusion, liver biomarkers at baseline and GVHD elsewhere are poor predictors of HGVHD

Alexander H. Yang, Mai Ai Thanda Han, and Niharika Samala contributed equally to this work.

Bisharah S. Rizvi and Rachel Marchalik contributed equally to this work.

on autopsy, and current clinical diagnostic criteria have unsatisfactory performance. Peak bilirubin and cholestatic injury predicted HGVHD on autopsy. A predictive model was developed accounting for changes over time. Further validation is needed.

INTRODUCTION

Allogeneic hematopoietic cell transplantation (HCT) is a potentially curative therapy for a variety of malignant and nonmalignant disorders. Hepatic complications are significant contributors to transplant-related morbidity and mortality.[1] The differential diagnosis for hepatic dysfunction in the posttransplant period includes graft-versus-host disease (GVHD), viral hepatitis, drug-induced liver injury (DILI), sepsis-associated cholestasis, immunotherapy-related hepatotoxicity, sinusoidal obstructive syndrome (SOS), malignancy, and many others.[2]

GVHD is one of the most commonly encountered complications of HCT and typically manifests in multiple organs, including the liver. It is a clinical diagnosis that is classified as acute GVHD (classic or late onset) and chronic GVHD (classic or overlap).[3] GVHD affects 30%–70% of allogeneic recipients, and the incidence varies with patient age, donor–recipient sex, differences in HCT regimens and graft sources, degree of human leukocyte antigen match, and donor relatedness.[4] Clinical suspicion for hepatic GVHD (HGVHD) diagnosis increases with an increase in alkaline phosphatase (ALP), alanine aminotransferase (ALT), and bilirubin; unfortunately, these are nonspecific markers of liver injury.[3,5] Furthermore, abnormal liver panels are associated with a wide range of disorders in the post-HCT population, especially in patients who are critically ill. On the other hand, accurate diagnosis is essential because misdiagnosis of HGVHD in the absence of other organ involvement can lead to inappropriate treatment with potent immunosuppressive agents, in turn increasing the risk of opportunistic infections, malignancy relapse, as well as other significant complications.[6]

There are currently no diagnostic clinical or histopathological criteria that can accurately diagnose and differentiate acute from chronic HGVHD, and final determination is dependent on the rest of the clinical picture as well as clinician judgment. Currently, the most widely used clinical diagnostic tool is the National Institutes of Health (NIH) Consensus Criteria (NCC) (previously the Keystone Criteria), with an elevated bilirubin and histologic finding of bile duct injury with mild lymphocytic infiltrate.[4] To our knowledge, no studies have been conducted assessing the agreement between clinical and histologic findings suggestive of HGVHD. This is likely because liver biopsies are not always possible to perform in the evaluation of post-HCT liver dysfunction due to the risk of bleeding and other related complications.[2] Moreover, exploring potential liver pathologies other than GVHD is necessary as liver injury in patients who are critically ill is often a likely clue to several simultaneous or interrelated processes.

In this study, we aim to characterize the patterns of liver injury seen in patients who are critically ill following HCT, assess the relationship between clinical and histologic findings in HGVHD, determine the immediate cause of death in order to determine both the accuracy of noninvasive diagnosis of HGVHD, and develop a predictive model for HGVHD on autopsy.

MATERIALS AND METHODS

Study design

This is a retrospective review of 51 patients at the NIH Clinical Center who were enrolled on an Institutional Review Board (IRB)-approved NIH protocol, underwent HCT between 2000 and 2012, and subsequently expired from various causes. Diagnosis of GVHD of various organs was recorded for biopsy (confirmed on biopsy, suspicious on biopsy, confirmed not on biopsy, and no biopsy), clinical diagnosis at time based on either Keystone Criteria (before 2005) or NCC per chart review (clinical diagnosis, clinical suspicion, and no clinical suspicion), and autopsy (confirmed on autopsy, not confirmed on autopsy). For baseline laboratory values, we recorded ALT, aspartate aminotransferase (AST), ALP, platelets, albumin, and bilirubin at 90 days before death. To account for the dynamic changes of tests, we recorded laboratory values up to 45 days before death in 15-day intervals: period −45, defined as 45 days to 31 days before death; period −30, defined as 30 days to 16 days before death; and period −15, defined as 15 days up until day of death or less if the death occurred in less than 45 days. We chose laboratory values with the least fluctuation to better represent the ongoing process of liver injury, including ALT, ALP, and bilirubin. We calculated the pattern of liver injury using R ratio (R) (outlined below) and determined the percentage of each pattern within each time interval, the average of R within each time interval (R mean), and the maximum bilirubin level. Values were recorded only

if there were more than two thirds of values available within the time interval.

Patients/legal guardians provided written informed consent to be followed under the IRB-approved protocol. All procedures were conducted in accordance with the IRB at the NIH.

Liver histology

An expert hepatopathologist (D.E.K.) retrospectively reviewed slides of livers from autopsies, batched and blinded. The slides were stained with hematoxylin and eosin, reticulin, and iron stains. Hepatic GVHD diagnosis was based on the NIH histopathologic consensus criteria as characterized by duct injury with absence of significant inflammation (mild lymphocytic infiltrate) after exclusion of other potential etiologies, mainly DILI.[4] Other pathologies were based on the following: nodular regenerative hyperplasia based on reticulin stain with findings of nodular areas of enlarged hepatocytes organized into two-cell-thick plates alternating with compressed liver cell plates; sepsis based on the presence of cholangiolar cholestasis; sinusoidal obstruction syndrome based on evidence of vascular narrowing in central veins; steatosis based on the presence of fat vacuoles in hepatocytes; steatohepatitis based on ballooning with or without Mallory bodies; and iron overload based on iron staining. Immediate and underlying causes of death were determined from the autopsy reports.

Characterization of the pattern of liver injury

The R, defined as serum ALT/upper limit of normal (ULN) divided by serum ALP/ULN, was calculated to determine the pattern of liver injury—hepatocellular, cholestatic, or mixed.[7] $R > 5$ defined a hepatocellular pattern, $R < 2$ a cholestatic pattern, and $2 < R < 5$ a mixed pattern of enzymes.[8,9]

Statistical analysis

Data were summarized using number and percentage, mean and SD, and median and interquartile range as appropriate. Chi-square tests and the Wilcoxon rank sum test were used to compare groups. Diagnosis accuracy was evaluated with the κ statistic. Logistic regression and mixed model linear regression with an unstructured covariance matrix were used to assess the relationships between laboratory data at −45 to −31, −30 to −16, and −15 to 0 days and HGVHD. All analyses were performed using SAS 9.4 (SAS Institute, Cary, NC).

RESULTS

Patient selection

A total of 51 patients who underwent an HCT and had autopsy results available at the NIH Clinical Center from 2000 to 2012 were reviewed. Sixteen (43%) were women with a median age of 49 years and a minimum age of 19 years. Six of these had received an autologous transplant and were excluded. Among the remaining 45 patients, 37 liver tissue samples from autopsies were adequate for the analysis of HGVHD. Eight samples were excluded because of the following: two specimens underwent autolysis and were difficult to interpret, two could not be located, three contained tumors (as this would affect local architecture and confound the results), and one was excluded by the pathologist due to fungal hyphae in the portal vein.

Characteristics of patients with autopsy liver tissue samples

Nine (24%) of the 37 samples had findings of bile duct injury with neutrophil infiltrate consistent with HGVHD on autopsy. Demographics and clinical characteristics of these groups are shown in Table 1. Only six had HGVHD as the only pathology present, the remaining samples showed evidence of iron overload (n = 8, 22%), hepatic steatosis (n = 7, 19%), nodular regenerative hyperplasia (NRH) (n = 3, 8%), and SOS (n = 3, 8%). Six autopsies had multiple pathologies in the liver (Table 2). Representative histologic images on autopsy are shown in Figure 1 (parenchymal atrophy, NRH, GVHD, veno-occlusive disease [VOD]-SOS, steatosis, and hemosiderosis).

Determination of liver injury pattern (R value calculation) and clinical diagnosis of HGVHD

At 90 days before death, we found there was no difference in AST, ALT, ALP, total bilirubin, direct bilirubin, platelet, or albumin between those with and without HGVHD (on autopsy). Nineteen (51%) patients presented with abnormal liver tests 90 days before death. Among the 51% of patients with an abnormal panel, 17 (89%) presented with a cholestatic pattern of liver injury while only two patients (11%) had a mixed pattern of liver injury. There were no patients with purely a hepatocellular pattern of liver injury (Table 1).

Twenty-five immediate causes of death were attributed to infection. Thirteen patients had clinical findings consistent with HGVHD based on Keystone Criteria or the NCC but did not show signs of HGVHD at autopsy. Eleven of those 13 patients died of

TABLE 1 Demographics, clinical characteristics, and laboratory data

| | | HGVHD on autopsy | | |
	Total (n = 37)	Yes (n = 9)	No (n = 28)	*p* value
Age at transplant (median, range)	49 (19–66)	49 (22–66)	49 (19–66)	0.56
Female sex	16 (43%)	3 (33%)	13 (46%)	0.70
Race				
White	25 (68%)	5 (56%)	20 (71%)	0.45
African American	7 (19%)	3 (33%)	4 (14%)	
Other	5 (14%)	1 (11%)	4 (14%)	
Underlying disease				
Hematologic malignancy	33 (89%)	8 (89%)	25 (89%)	
Nonhematologic malignancy	3 (8%)	1 (11%)	2 (7%)	
Immune errors of immunity	1 (3%)	0	1 (4%)	
Type of transplant				
Matched, related, peripheral	16 (43%)	6 (67%)	10 (36%)	
Matched, unrelated, peripheral	8 (22%)	2 (22%)	6 (21%)	
Mismatch, unknown relation	2 (5%)	0	2 (7%)	
Cord blood	3 (8.1%)	0	3 (10.7%)	
NA	8 (21.6%)	1 (11%)	7 (25%)	
Conditioning				
Fludarabine/cyclophosphamide	24 (64.9%)	5 (56%)[a]	19 (67.9%)	
Pentostatin/cyclophosphamide	2 (5%)	1 (11%)[b]	1 (4%)	
Larson protocol	1 (3%)	1 (11%)	0	
Pentostatin + dexamethasone	1 (3%)	0	1 (4%)	
NA	9 (24.3%)	2 (22%)	7 (25%)	
Immunosuppression				
Cyclosporin only	11 (30%)	3 (33%)	8 (29%)	
Sirolimus only	2 (5%)	1 (11%)	1 (4%)	
Sirolimus + cyclosporin	10 (27%)	2 (22%)	8 (28.6%)[c]	
Sirolimus + tacrolimus	5 (14%)	0	5 (18%)	
Cyclosporin + methotrexate	3 (8%)	1 (11%)	2 (7%)	
Tacrolimus + sirolimus + methotrexate	1 (3%)	0	1 (4%)	
Alemtuzumab + cyclosporin	4 (11%)	1 (11%)	3 (11%)	
NA	1 (2.7%)	1 (11%)	0 (0%)	
Time between transplant and death (days)	147 (77–283)	283 (140–303)	140 (75–215)	0.10
Second transplant	9 (23%)	2 (22%)	7 (18%)	
Donor lymphocyte infusion	8 (21.6%)	2 (22.2%)	7 (25%)	
Number of organs with GVHD before death (median, range)				
Biopsy confirmed	1 (0–3)	1 (0–2)	1 (0–3)	
Biopsy suspicion	0 (0–1)	0 (0–1)	0 (0–1)	
Clinical confirmed	2 (0–5)	2 (0–5)	2 (0–5)	
Clinical suspicion	0 (0–2)	0 (0–1)	0 (0–2)	
Autopsy confirmed	1 (0–2)	1 (1–2)	0 (0–2)	
Autopsy suspicion	0 (0–1)	0 (0–1)	0 (0–1)	
Laboratory variables at 90 days before death				
Alkaline phosphatase (U/L)	112 (84–236)	148 (96–247)	111 (79–226)	0.59

(Continues)

TABLE 1 (Continued)

| | | HGVHD on autopsy | | |
	Total (n = 37)	Yes (n = 9)	No (n = 28)	*p* value
Alanine aminotransferase (U/L)	26 (18–48)	29 (22–39)	26 (18–50)	0.59
Total bilirubin (mg/dl)	0.6 (0.4–0.9)	0.9 (0.5–2.1)	0.6 (0.4–0.9)	0.21
Direct bilirubin (mg/dl)	0.2 (0.1–0.5)	0.4 (0.1–1.6)	0.2 (0.1–0.4)	0.58
Platelet (10^3/μl)	52 (28–90)	75 (34–90)	48 (25–109)	0.24
Albumin (g/dl)	2.7 (2.4–3.3)	2.9 (2.3–3.3)	2.7 (2.5–3.3)	0.90
Liver injury pattern at 90 days before death				
Normal ALP and ALT	18 (49%)	4 (44%)	14 (50%)	0.68
Cholestatic	17 (50%)	4 (44%)	13 (46%)	
Hepatocellular	0	0	0	
Mixed	2 (5%)	1 (11%)	1 (4%)	

Note: Data presented as median (interquartile range) or n (%) unless otherwise indicated.

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; HGVHD, hepatic graft-versus-host disease; NA, not available.

[a]Two patients with reduced intensity.

[b]Reduced intensity.

[c]One patient received sirolimus for 3 days then switched to cyclosporin.

TABLE 2 Histologic distribution of liver injury

|	Total	Clinically suspected HGVHD	Not clinically suspected HGVHD
No obvious liver disease	14	6	8
GVHD alone	6	0	6
GVHD + NRH + steatosis	1	0	1
GVHD + SOS	1	1	0
GVHD + steatosis + iron overload	1	1	0
NRH alone	1	1	0
NRH + SOS + steatosis	1	0	1
Steatosis alone	2	2	0
Steatosis + iron overload	2	1	1
SOS alone	1	0	1
Iron overload alone	5	2	3
Others (ischemic hepatitis)	2	1	1
Total	37	15	22

Abbreviations: GVHD, graft-versus-host disease; HGVHD, hepatic graft-versus-host disease; NRH, nodular regenerative hyperplasia; SOS, sinusoidal obstructive syndrome.

infection-related causes. Underlying patient diagnoses are also listed (Table S1).

Association between GVHD before death and HGVHD on autopsy

When evaluating whether biopsy or clinical diagnosis of GVHD had any association with HGVHD on autopsy, we found that there were no associations using Fisher's exact test. Table S2).

To compare the time between diagnosis of GVHD and death, we compared those with and without GVHD on autopsy and time of biopsy-proven GVHD. There were two patients (22.2%) with no organ involvement before death that shows HGVHD on autopsy compared to six (21.4%) with no HGVHD on autopsy; there were five patients (55.6%) with a median time of 38 days (range 7–213 days) with one biopsy-proven organ involvement in those with HGVHD on autopsy compared to fourteen (50%) with a median time of 35 days (range 2–2463 days) with
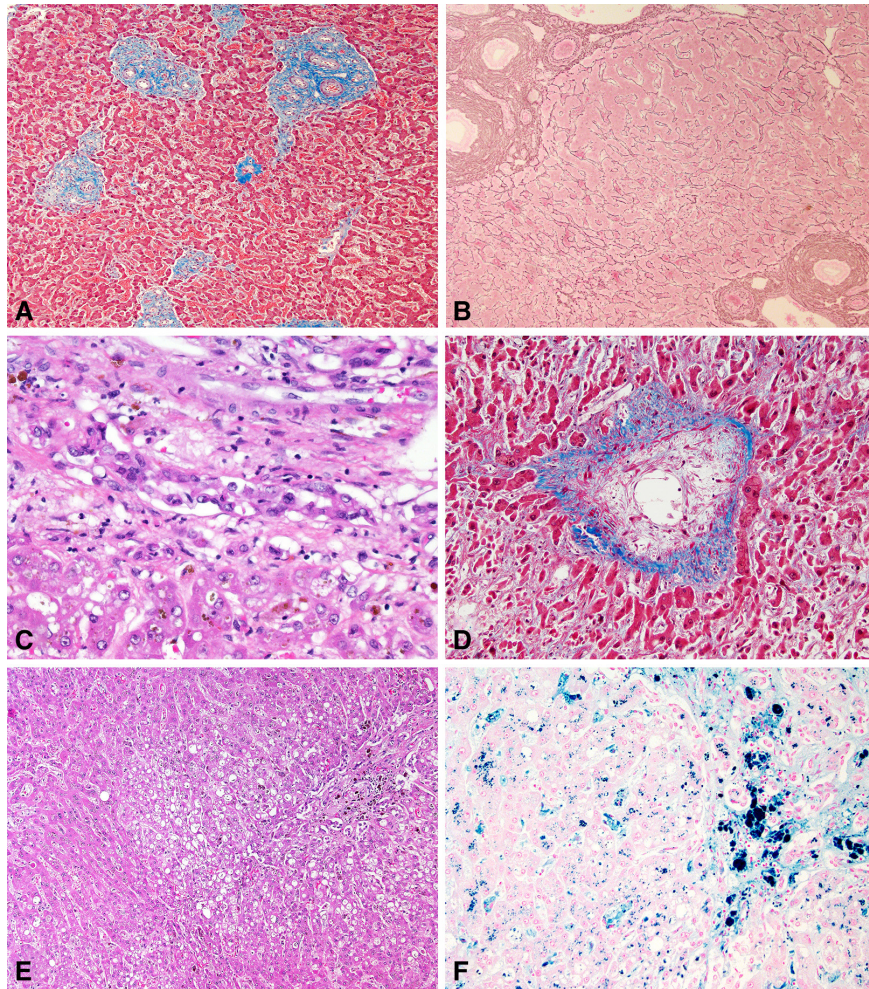
FIGURE 1 Liver histological findings at autopsy after hematopoietic cell transplantation. (A) Parenchymal atrophy. The portal areas are closely spaced and lack portal veins (Masson trichrome, ×100). (B) Nodular regenerative hyperplasia. There is a nodular region with wide liver cell plates bounded by atrophic, narrowed, liver cell plates (reticulin, ×100). (C) Graft-versus-host disease. Bile duct injury with enlarged reactive epithelial nuclei and several infiltrating lymphocytes (H&E, ×400). (D) Veno-occlusive disease–sinusoidal obstructive syndrome. This large central vein shows narrowing by loose connective tissue (Masson trichrome, ×200). (E) Steatosis. Moderate macrovesicular steatosis in a zone 1 distribution (H&E, ×100). (F) Hemosiderosis. Moderate to marked iron accumulation in hepatocytes and reticuloendothelial cells (Iron, ×200). Abbreviation: H&E, hematoxylin and eosin

one biopsy-proven organ involvement in those without HGVHD on autopsy. In those with two or more biopsy-proven organ involvement, there were two (22.2%) with a time of 13 and 2157 days in those with HGVHD on autopsy compared to eight (28.6%) with a median time of 20.5 days (range 5–188 days) in those without HGVHD on autopsy.

Four patients had liver biopsies before death; one patient had findings suggestive of sepsis versus HGVHD on biopsy and had HGVHD on autopsy, with 7 days between the two samples. One had neither HGVHD on biopsy nor autopsy, with 7 days between the two samples. Two had HGVHD on biopsy but none on autopsy, with 45 and 86 days between biopsy and death.

To understand the discordance between clinical suspicion of GVHD and autopsy findings, we considered the following two scenarios: 16 patients had clinical suspicion for liver GVHD, 12 of those had no GVHD on autopsy, and only one had no other clinical suspicion or biopsy-confirmed GVHD in other organs. The patient was a 31-year-old man with a history of Epstein-Barr virus (EBV) lymphoproliferative disorder who underwent a second transplant and had a liver biopsy on day +6 for increasing bilirubin and liver enzymes, with concern for VOD versus GVHD. The biopsy showed EBV-positive large B-cell lymphoma, with adjacent liver showing intrahepatic cholestasis, erythrophagocytosis, and Mallory bodies. The patient died on day +29, and cause of death was multiorgan failure from EBV-associated lymphoproliferative disorder involving lungs, liver, right kidney, and mediastinal lymph nodes.

In the second scenario, 18 patients did not have clinical suspicion for liver GVHD; five of those had GVHD on autopsy, and two of those had no other clinical suspicion or biopsy confirmed GVHD in other organs. The first patient, a 50-year-old woman with relapsed,

diffuse, large B-cell lymphoma, died on day +234 from pulmonary edema with diffuse alveolar damage and was found to have evidence of HGVHD on autopsy. The second patient was a 58-year-old woman with refractory, recurrent, diffuse, large B-cell lymphoma who had a second transplant and died on day +166 from polymicrobial pneumonia from aspiration and was found to have evidence of HGVHD on autopsy.

Assessment of the agreement between clinical and histopathologic findings

Fifteen out of 37 patients with liver autopsies showed clinical diagnosis of HGVHD at the same time. Of those, HGVHD was histologically confirmed in only two cases (13%), while others had histologic characteristics of NRH, SOS, iron overload, or steatosis. Twenty-two of the 37 patients showed no clinical signs of HGVHD. Of those, seven (32%) had histologic evidence suggestive of HGVHD despite the lack of clinical findings consistent with GVHD. Sensitivity, specificity, positive predictive value, and negative predictive value for clinical diagnosis predicting histopathologic evidence of HGVHD at autopsy using abnormal liver tests alone were 22%, 54%, 13%, and 68%, respectively. No correlation was noted between the clinical and histologic findings characteristic of HGVHD ($\kappa = -0.20$; 95% confidence interval [CI], -0.47 to -0.08). Fourteen percent of the patients had multiple pathologies on liver histology (Table 2). Of note, a total of 17 patients (11 among biopsied non-HGVHD and six among biopsied-HGVHD) received systemic immunosuppression during the period in which laboratories were drawn.

Association between dynamic markers and HGVHD on autopsy

To evaluate laboratory markers (mean R, maximum bilirubin) and liver injury pattern (hepatocellular pattern and cholestatic), spaghetti plots of values (mean R, maximum bilirubin, percentage of cholestatic pattern, and percentage of hepatocellular pattern) were shown at three time periods (Figure S1).

To compare markers across time, we included those patients with data points at all three time periods (six with HGVHD and 16 with non-HGVHD on autopsy) and used a mixed linear model. We found that those with HGVHD on autopsy had a higher peak bilirubin at period −15 (18.5 vs. 6.9 mg/dl, $p = 0.011$). When comparing across time, those with HGVHD had an increase in peak bilirubin from period −45 to period −15 (9.0 vs. 18.5 mg/dl, $p = 0.018$) and period −30 to period −15 (10.9 vs. 18.5 mg/dl, $p = 0.013$). There were no differences in the non-HGVHD group when compared across time (Figure S2A). For mean R, there were no differences between those with HGVHD and non-HGVHD on autopsy and no differences within groups across time points (Figure S2B). For percentage of cholestatic injury, those with HGVHD on autopsy had a higher percentage of cholestatic injury at period −15 (82% vs. 45%, $p = 0.032$), and there were no differences within groups across time points (Figure S2C). For percentage of hepatocellular injury, there were no differences between those with HGVHD and non-GVHD on autopsy and no differences within groups across time points (Figure S2D). Analysis including all data while accounting for individuals with multiple data is shown in Figure S3, and details of the analyses are show in Tables S3–S6.

Predictive modeling of HGVHD on autopsy

Using the above findings, we performed logistic regression using peak bilirubin and percentage of cholestatic injury at each time period to develop a predictive model. Of the models shown in Table S7, using the peak bilirubin at period −15 resulted in an area under the receiver operating characteristic curve (AUROC) of 0.80 (95% CI, 0.62–0.98), corresponding to a bilirubin level of 10.12 mg/dl. Adding on the percentage of cholestatic injury at period −15 resulted in an AUROC of 0.81 (95% CI, 0.65–0.97); and when adding on a factor of the difference of peak bilirubin at period −30 and period −15, an AUROC of 0.77 (95% CI, 0.52–1.00) was achieved (Figure 2). Performance of these models with cut-off values are shown in Tables S8 and S9.

DISCUSSION

Hepatic dysfunction can contribute considerably to substantial morbidity and mortality after allogeneic HCT, accounting for mortality in more than 30% of patients.[10] The risk factors predicting liver disease in post-HCT include advanced age, use of an alternative donor source, medications given pretransplant (e.g., inotuzumab or gemtuzumab), transplant regimen (e.g., busulfan, cyclophosphamide, sirolimus), and preexisting liver disease. The diagnosis of liver dysfunction is often clinical as these patients are at a higher risk for biopsy-related complications secondary to thrombocytopenia or coagulopathy.[2,11] In fact, a study reviewing the safety of percutaneous liver biopsies using Klatskin needles in 3357 biopsies observed that patients with a diagnosis of NRH (odds ratio [OR], 17), DILI (OR, 20), GVHD (OR, 32), and hepatocellular carcinoma (OR, 34) are far more likely to suffer a major postprocedure complication.[12] Our data suggest that clinical findings and liver biomarkers may not be sensitive or specific diagnostic tools for liver injury after allogeneic HCT in patients who are critically ill as multiple causes could cause laboratory test abnormalities. Because hepatic
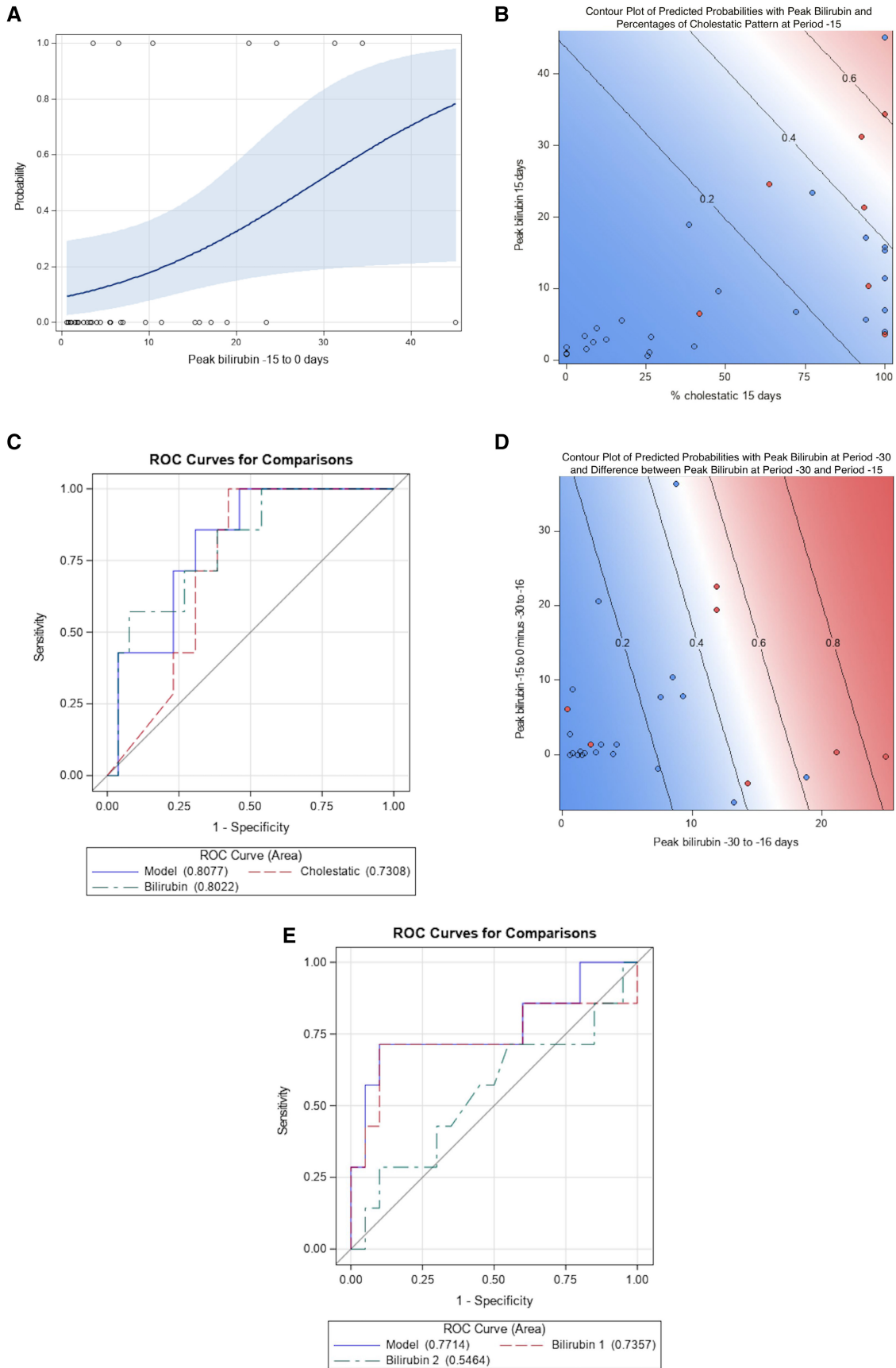
FIGURE 2 Regression model. (A) Contour curve for peak bilirubin at period −15. (B) Contour plot with peak bilirubin and percentage of cholestatic pattern at Period −15. (C) Receiving operator curve (ROC) with peak bilirubin, percentage of cholestatic pattern and model at period −15. (D) Contour plot with peak bilirubin at period −30 and difference between peak bilirubin at period −30 and period -15. (E) ROC with peak bilirubin at period −30 and difference between peak bilirubin at period -30 and period −15

dysfunction strongly correlates with intensive care unit mortality, it is essential to determine the correct diagnosis underlying clinical liver dysfunction in order to initiate appropriate therapy.[13]

The majority (n = 24, 53%) of our patients presented with an abnormal liver panel 90 days before death. Of those patients, 21 (88%) had a cholestatic pattern and three (12%) were mixed. These percentages are comparable to those seen in other studies. For instance, Kim et al.[14] observed liver dysfunction during the first year post-HCT in greater than 80% of allogeneic bone marrow transplant recipients. In another study by Farthing et al.,[15] hepatic panel abnormalities occurred in 83% of patients and the severity of liver test abnormalities was associated with increased mortality. The appearance of a cholestatic pattern is not uncommon after HCT and is seen in many common hepatic complications of HCT (infection, GVHD, DILI, sepsis-associated cholestasis, disease relapse, and SOS).[1,2,16]

Typically, GVHD, DILI, sepsis, and SOS are the most common causes of post-HCT hepatic dysfunction.[2,14] However, the highest morbidity and mortality are seen in severe cases of adenovirus hepatitis, fungal infection of the liver, SOS, and GVHD.[17] In our cohort, the most commonly seen histologic features were associated with HGVHD (n = 9, 24%), hepatic steatosis (n = 7, 19%), and iron overload (n = 8, 22%). Less common findings were SOS, NRH, and ischemic hepatitis; six samples showed mixed pathologies. Of note, a significant portion of the cohort (n = 14, 38%) had no microscopic evidence of liver injury. Although only three patients presented with SOS (either alone or overlapping with other pathologies), it is peculiar to see SOS in the long-term post-HCT course as SOS typically occurs within the first 21 days following HCT (in patients that have received myeloablative therapy).[18]

Twenty-five of the 37 patients (67.6%) in the autopsy cohort ultimately died of infection-related causes. Of the 15 patients who satisfied the clinical criteria for HGVHD, 13 (87%) did not show histopathological signs of disease. Almost all (n = 12, 80%) died of infection-related causes. These results are alarming and reveal the danger of treating GVHD without an established diagnosis. In our cohort, only one patient had clinical suspicion for HGHVD with no other GVHD but showed no HGVHD on autopsy; biopsy of liver during life showed no evidence of GVHD. The cause of death was multiorgan failure from EBV-associated lymphoproliferative disorder. There are many consequences of GVHD overtreatment, including administering unnecessarily toxic therapies, infection, loss/impairment of graft-versus tumor immunity, as well as failure to diagnose and appropriately manage other underlying disorders.[4,6]

We found that there was no association between clinical and biopsy GVHD of gastrointestinal, skin, and liver organs before death and HGVHD on autopsy (Table S2). The median time between biopsy-proven GVHD of one organ and death was 38 days in those with HGVHD compared to 35 days in those with no HGVHD on autopsy; for those with two or more biopsy-proven organ involvement, the median time was 1086 days in those with HGVHD on autopsy compared to 20.5 days in those without HGVHD on autopsy. When looking at laboratory values at 90 days before death, no statistical differences in liver tests or pattern of liver injury (Table 1) could be detected in patients with or without histologic features of HGVHD, suggesting that these laboratory values do not guide diagnosis in this cohort. As no agreement was found between clinical and histologic findings characteristic of HGVHD (κ = −0.20; 95% CI, −0.47 to −0.08), it appears that abnormal liver tests alone are not sufficient to diagnose HGVHD in this cohort accurately. Unfortunately, other than liver biopsy, there are currently no clear guidelines or tools that could serve as a better diagnostic tool for HGVHD.[4,19]

To account for the dynamic changes of liver injury, we developed a method to compare laboratory parameters, including peak bilirubin, mean R, percentage cholestatic, and percentage hepatocellular at three time periods: period −45 (days 31 to 45 before death), period −30 (days −16 to 30 before death), and period −15 (days −15 up to death). We found that those with HGVHD on autopsy had a higher peak bilirubin at period −15 compared to those without HGVHD on autopsy and an increase over time. In terms of mean R, we noted that those with HGVHD on autopsy had a higher mean R at period −30. In terms of injury pattern, those with HGVHD on autopsy had a higher percentage of cholestatic pattern at period −15 and an increase from period −30 (including all data); there was no difference in hepatocellular injury.

With this finding of dynamic changes in laboratory values, we developed logistic regression to predict HGVHD on autopsy and found that a combination of peak bilirubin at period −15 with either percentage of cholestatic injury pattern at period −15 or difference of peak bilirubin at period −30 and period −15 was able to predict HGVHD on autopsy (Table S5). Using a univariate model of peak bilirubin at period −15 corresponds to a cutoff of 10.12 mg/dl, which is higher than the NCC. Two other models combining either percentage of cholestatic injury at period −15 and increase of peak bilirubin from period −30 to −15 also had good performance, suggesting further improvement of the NCC. While this is a retrospective study looking at premortem laboratories, this finding is valuable in that if an increase in bilirubin and change toward a cholestatic injury pattern is found during the posttransplant course, the suspicion for GVHD would be increased.

Limitations of this study include a retrospective design that can introduce selection and information bias. Temporal relationships can also be challenging to assess due to the retrospective nature of the study and the lack of serial sampling of the liver. As the patients are typically ill and may receive platelet or albumin

transfusions, it is difficult to compare the values between those with and without GVHD. Additionally, as the majority of patients were taking immunosuppressant medications during the evaluation, there is the possibility that histopathological features of GVHD were obscured or treated before autopsy sampling. However, this study provides new information on the agreement between clinical and histologic findings suggestive of HGVHD, with potential important implications for developing better patient management strategies.

In summary, patients who are critically ill following HCT will commonly present with abnormal liver tests, typically with a cholestatic pattern. Considering broad differential diagnoses in addition to HGVHD remains essential in this population. Because static abnormal liver tests alone are not sufficient to diagnose all cases of HGVHD, we accounted for the dynamic changes that added to the diagnostic accuracy of HGVHD on autopsy. In addition, biopsy should be considered despite the risks and is especially important in cases where the liver injury progresses despite appropriate treatment for the presumed clinical diagnosis. Further, caution should be exercised when deciding whether to initiate or escalate treatment as potent immunosuppressive agents can increase the risk of opportunistic infections and organ toxicity, leading to increased morbidity and mortality. Future prospective multidisciplinary studies are an imperative to further evaluate the complex nature of liver injury in the critically ill postallogeneic HCT population.

CONFLICT OF INTEREST
Nothing to report.

AUTHOR CONTRIBUTIONS
All authors contributed to the writing and editing of the manuscript. Data collection: Mai Ai Thanda Han, Niharika Samala, Bisharah S. Rizvi, Ohad Etzion, Ruchi Patel, Vinshi Khan, Devika Kapuria, Vikramaditya Samala Venkat. Significant editing: Rachel Marchalik. Data Collection, significant editing, finalization of manuscript: Alexander H. Yang. Statistical analysis: Elizabeth C. Wright. Expert hepatopathologist, reviewed all autopsy slides: David E. Kleiner. Data analysis, intellectual input: Christopher G. Kanakry. Care of patients, intellectual input: Jennifer A. Kanakry, Christopher G. Kanakry, Steven Pavletic, Kirsten M. Williams. Intellectual input, supervision, patient care, editing: Theo Heller.

ORCID
Alexander H. Yang https://orcid.org/0000-0001-5975-5131

REFERENCES
1. Sakai M, Strasser SI, Shulman HM, McDonald SJ, Schoch HG, McDonald GB. Severe hepatocellular injury after hematopoietic cell transplant: incidence, etiology and outcome. Bone Marrow Transplant. 2009;44:441–7.
2. Matsukuma KE, Wei D, Sun K, Ramsamooj R, Chen M. Diagnosis and differential diagnosis of hepatic graft versus host disease (GVHD). J Gastrointest Oncol. 2016;7(Suppl 1): S21–31.
3. Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW, 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;21:389–401.e1.
4. Shulman HM, Cardona DM, Greenson JK, Hingorani S, Horn T, Huber E, et al. NIH Consensus development project on criteria for clinical trials in chronic graft-versus-host disease: II. The 2014 Pathology Working Group Report. Biol Blood Marrow Transplant. 2015;21:589–603.
5. Lee SJ. Classification systems for chronic graft-versus-host disease. Blood. 2017;129:30–7.
6. Jacobsohn DA, Montross S, Anders V, Vogelsang GB. Clinical importance of confirming or excluding the diagnosis of chronic graft-versus-host disease. Bone Marrow Transplant. 2001;28:1047–51.
7. Chalasani NP, Hayashi PH, Bonkovsky HL, Navarro VJ, Lee WM, Fontana RJ, et al. ACG Clinical Guideline: the diagnosis and management of idiosyncratic drug-induced liver injury. Am J Gastroenterol. 2014;109:950–66.
8. Benichou C, Danan G, Flahault A. Causality assessment of adverse reactions to drugs--II. An original model for validation of drug causality assessment methods: case reports with positive rechallenge. J Clin Epidemiol. 1993;46:1331–6.
9. Danan G, Benichou C. Causality assessment of adverse reactions to drugs--I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. J Clin Epidemiol. 1993;46:1323–30.
10. El-Sayed MH, El-Haddad A, Fahmy OA, Salama II, Mahmoud HK. Liver disease is a major cause of mortality following allogeneic bone-marrow transplantation. Eur J Gastroenterol Hepatol. 2004;16:1347–54.
11. Levitsky J, Sorrell MF. Hepatic complications of hematopoietic cell transplantation. Curr Gastroenterol Rep. 2007;9:60–5.
12. Takyar V, Etzion O, Heller T, Kleiner DE, Rotman Y, Ghany MG, et al. Complications of percutaneous liver biopsy with Klatskin needles: a 36-year single-centre experience. Aliment Pharmacol Ther. 2017;45:744–53.
13. Damm TW, Kramer DJ. The liver in critical illness. Crit Care Clin. 2016;32:425–38.
14. Kim BK, Chung KW, Sun HS, Suh JG, Min WS, Kang CS, et al. Liver disease during the first post-transplant year in bone marrow transplantation recipients: retrospective study. Bone Marrow Transplant. 2000;26:193–7.
15. Farthing MJ, Clark ML, Sloane JP, Powles RL, McElwain TJ. Liver disease after bone marrow transplantation. Gut. 1982;23:465–74.
16. Hogan WJ, Maris M, Storer B, Sandmaier BM, Maloney DG, Schoch HG, et al. Hepatic injury after nonmyeloablative conditioning followed by allogeneic hematopoietic cell transplantation: a study of 193 patients. Blood. 2004;103:78–84.
17. Tuncer HH, Rana N, Milani C, Darko A, Al-Homsi SA. Gastrointestinal and hepatic complications of hematopoietic stem cell transplantation. World J Gastroenterol. 2012;18: 1851–60.
18. McDonald GB. Hepatobiliary complications of hematopoietic cell transplantation, 40 years on. Hepatology. 2010;51: 1450–60.

19. Dignan FL, Amrolia P, Clark A, Cornish J, Jackson G, Mahendra P, et al. Haemato-oncology Task Force of British Committee for Standards in Haematology; British Society for Blood and Marrow Transplantation. Diagnosis and management of chronic graft-versus-host disease. Br J Haematol. 2012;158:46–61.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.