Significant morbidity and mortality have been associated with liver complications after allogeneic hematopoietic stem cell transplantation (allo-HSCT). Causes and consequences of these hepato-biliary complications are various and might be life-threatening. A high misdiagnosis rate has been reported because of a weak correlation between clinical, laboratory and imaging data. Liver elastography, a liver stiffness measure, is able to assess liver fibrosis and portal hypertension in most liver diseases, but data after allo-HSCT are scarce. Our aim was to determine the interest of sequential liver stiffness measurements for the diagnosis of early hepatic complications after allo-HSCT. Over a 2-year time period, 161 consecutive adult patients were included and 146 were analyzed. Ultrasonography and elastography measurements were performed before transplantation, at day+7 and day+14 by three different experienced radiologists unaware of the patients’ clinical status. Eighty-one (55%) patients had liver involvements within the first 100 days after allo-HSCT. Baseline elastography was not predictive for the occurrence of overall liver abnormalities. A significant increase in two-dimensional real-time shearwave elastography (2D-SWE) was found in patients with sinusoidal obstruction syndrome (SOS). Fifteen patients (10%) fulfilled European Society for Blood and Marrow Transplantation (EBMT) score criteria and twelve (8%) reached Baltimore criteria for SOS diagnosis, but only six (4%) had a confirmed SOS. 2D-SWE at day+14 allowed early detection of SOS (AUROC=0.84, P=0.004) and improved sensitivity (75%), specificity (99%) and positive predictive value (60%) over the Seattle, Baltimore or EBMT scores. A 2D-SWE measurement above 8.1 kPa at day+14 after allo-HSCT seems a promising, non-invasive, and reproducible tool for early and accurate diagnosis of SOS.

Elastography improves accuracy of early hepato-biliary complications diagnosis after allogeneic stem cell transplantation


1Hematology and Transplantation Unit, Saint Louis Hospital, APHP, Paris; 2Radiology Unit, Saint Louis Hospital, APHP, Paris; 3DHU Unit, Pôle des Maladies de l’Appareil Digestif, Service d’Hépatologie, Centre de Référence des Maladies Vasculaires du Foie, Hôpital Beaujon, AP-HP, Clichy and 4Université de Paris, INSERM U976, Paris, France

*PED and PB contributed equally as co-first authors.
#GS and DM contributed equally as co-senior authors

Absract

Significant morbidity and mortality have been associated with liver complications after allogeneic hematopoietic stem cell transplantation (allo-HSCT). Causes and consequences of these hepato-biliary complications are various and might be life-threatening. A high misdiagnosis rate has been reported because of a weak correlation between clinical, laboratory and imaging data. Liver elastography, a liver stiffness measure, is able to assess liver fibrosis and portal hypertension in most liver diseases, but data after allo-HSCT are scarce. Our aim was to determine the interest of sequential liver stiffness measurements for the diagnosis of early hepatic complications after allo-HSCT. Over a 2-year time period, 161 consecutive adult patients were included and 146 were analyzed. Ultrasonography and elastography measurements were performed before transplantation, at day+7 and day+14 by three different experienced radiologists unaware of the patients’ clinical status. Eighty-one (55%) patients had liver involvements within the first 100 days after allo-HSCT. Baseline elastography was not predictive for the occurrence of overall liver abnormalities. A significant increase in two-dimensional real-time shearwave elastography (2D-SWE) was found in patients with sinusoidal obstruction syndrome (SOS). Fifteen patients (10%) fulfilled European Society for Blood and Marrow Transplantation (EBMT) score criteria and twelve (8%) reached Baltimore criteria for SOS diagnosis, but only six (4%) had a confirmed SOS. 2D-SWE at day+14 allowed early detection of SOS (AUROC=0.84, P=0.004) and improved sensitivity (75%), specificity (99%) and positive predictive value (60%) over the Seattle, Baltimore or EBMT scores. A 2D-SWE measurement above 8.1 kPa at day+14 after allo-HSCT seems a promising, non-invasive, and reproducible tool for early and accurate diagnosis of SOS.

Introduction

Over the past two decades, overall survival rate after allo-HSCT has improved. However, transplantation-related mortality (TRM) remains a significant cause of death, with a reported 15% to 35% rate in the current era. Hepato-biliary complications lead to significant morbidity and TRM after allo-HSCT. They include liver graft-versus-host disease (GVHD), sinusoidal obstruction syndrome (SOS), drug-induced hepatotoxicity, cholangitis lenta, malignant infiltration, iron overload, hemodynamic modification and biliary obstruction. Previous studies have reported a higher TRM rate in patients who had a high level of bilirubinemia.
(above 4 mg/dL or 68 μmol/L) after allo-HSCT. More recently, even after reduced-intensity conditioning (RIC), 20% to 26% of patients were reported to have a bilirubin peak above 4 mg/dL, which was also associated with poor outcomes. Clinical, biological and imaging data are weekly correlated and can lead to delayed or wrong diagnoses. Although liver biopsy is useful to establish diagnosis, it has limitations including its feasibility in severely-ill patients and bleeding risk in case of severe thrombocytopenia. Several clinical scores have been developed to diagnose SOS after allo-HSCT and currently lack of specificity and sensitivity. We previously observed that among patients who fulfilled clinical criteria for SOS diagnosis, less than a half had a final diagnosis of SOS after liver biopsy. Pitfall associated with ultrasonography and Doppler are due to heterogeneity and lack of reproducibility and to the late onset of some radiological signs (e.g., reverse flow in the portal vein) that can occur in SOS. There are currently no early ultrasonography specific signs that are able to discriminate these various hepatic complications. Liver stiffness measurement using elastography is widely used and recommended for the assessment of liver fibrosis, cirrhosis and portal hypertension. Different techniques have been described: transient elastography (TE with FibroScan), point shear wave with acoustic radiation force impulse (ARFI), and two-dimensional real-time shear wave (2D-SWE). The objective of this study was to determine the feasibility and interest of sequential measures of liver elastography for the diagnosis of early hepatic complications after allo-HSCT.

**Methods**

**Patients**

Between July 2017 and July 2019, 212 patients underwent an allo-HSCT in the Saint Louis Hospital (Paris, France). A total of 161 patients were included. This study has been conducted in compliance with the Declaration of Helsinki. All patients gave their written consent for the registration of clinical and biological data (CNIL number 2093819), were collected and processed anonymously in a dedicated study (CNIL number 2211540), with authorization of the IRB 0000388 (study number 20-697).

**Ultrasonography and elastography**

Ultrasonography, Doppler, and elastography were performed at baseline, at day+7, and at day+14. Two methods were used for elastography for all patients: transient elastography with Fibroscan® (Echosens, Paris, France) and 2D-SWE (Axiplorer, SuperSonic Imaging SA, Aix-en-Provence, France) with a 3.5 MHz convex ultrasound probe (SCX-6-1) and a 7.5 MHz linear ultrasound probe (SL-10-2). For all ultrasonography and Doppler examinations, the following criteria were assessed: liver and splenic measurements, measurement of the gallbladder wall, ascites, portal vein diameter, portal vein direction flow and maximal flow velocity, spectral waveforms of the hepatic veins. Based on Lassau et al., and European Society for Blood and Marrow Transplantation (EBMT) classification, an ultrasonod-Doppler score based on seven criteria was performed: (i) hepatomegaly, (ii) splenomegaly, (iii) gall bladder wall thickening, (iv) dilatation of main portal vein, (v) ascites, (vi) decrease mean velocity of portal vein, (vii) hepatofugal flow or no flow in portal vein. An additional ultrasonography and elastography could be performed at the discretion of the physician.

**Liver test and definition of liver involvements**

Liver involvement was considered if increased serum aspartate aminotransferase (AST) or alanine aminotransferase (ALT) level above twice the upper limit of normal values or hyperbilirubinemia (above 17 μmol/L) occurred in two consecutive measures. All medical records were retrospectively reviewed to determine the final liver diagnosis. GvHD was graded according to the modified Glucksberg’s classification. When patients had no other organ involvement than suspected liver GvHD, a biopsy was performed to ascertain the diagnosis. SOS diagnosis was suspected when EBMT, Baltimore or modified Seattle clinical criteria were present in patients. Diagnosis was retained only if proven on liver biopsy, or using ultrasonography and Doppler criteria, as described in EBMT classification. If not proven on biopsy or ultrasonography, SOS diagnosis was considered only in the absence of infectious disease, drug toxicity or GvHD, as recommended by the European Association for the Study of the Liver (EASL) guidelines. Drug-induced liver injury (DILI) was defined according to EASL guidelines.

**Statistical analysis**

Two-group comparisons were performed with Mann-Whitney U test and multiple comparisons were performed with Kruskal-Wallis test followed by a Dunn’s correction for multiple comparisons. Two-way ANOVA test followed by Dunnet correction was used for multiple comparisons of data with normal distribution and equal variance. ROC curves were built for the assessment of liver fibrosis, cirrhosis and portal hypertension with liver biopsy (n=3), or using ultrasonography and Doppler criteria, as described in EBMT classification. If not proven on biopsy or ultrasonography, SOS diagnosis was considered only in the absence of infectious disease, drug toxicity or GvHD (n=3), as recommended by the European Association for the Study of the Liver (EASL) guidelines. All statistical tests were two-tailed with a significance level of 0.05.

**Results**

**Population overview**

Over two years, 146 patients out of 161 consecutive patients were analyzed (Figure 1). Fifteen patients were excluded for analysis due to incomplete ultrasonography evaluations. Main patient, disease and transplant features are summarized in Table 1 and Table 2. Median follow-up was 9.2 months (range, 3-19). Five donors had prior hepatitis B virus (HBV) hepatitis, but none had viral replication at the time of stem cell collection. None of the patients had detectable HBV DNA or hepatitis C virus (HCV) RNA. Two patients had Child-Pugh A cirrhosis (previous HBV infection and telomeropathy, one each). Eighteen (12%) and seven (5%) patients had prior cholecystectomy or splenectomy, respectively and could not be evaluated for all the ultrasonography criteria. During follow-up, 82 patients (22%) died, including 16 (11%) early deaths (before day+100). The leading cause of early death was TRM (93%), including three patients with SOS and two with liver GvHD who had all received a RIC regimen.

**Incidence of hepatic involvements after allogeneic hematopoietic stem cell transplantation**

Eighty-one (55%) patients had a hepatic involvement defined by an elevation of liver enzymes and/or hyperbilirubinemia during the first 100 days after allo-HSCT (Figure 2). Hepatic GvHD was diagnosed in 11 patients...
(including three cases that were proven after liver biopsy) and were associated with skin or gut GvHD in all cases but one. Liver tests improved after GvHD treatment in six cases. SOS diagnosis was suspected when patients had clinical criteria based on classifications (EBMT, Seattle and Baltimore). SOS diagnosis was retrospectively retained when proven on liver biopsy (n=3) or in patients with usual ultrasonography and Doppler criteria (n=3), in the absence of other causes (GvHD, infectious or drug-related) and according to the clinical evolution. Drug-induced liver injury was observed in 48 patients after exposure to drugs within the conditioning regimen (n=52), azole antifungal therapy (n=12) or cyclosporine A (n=4), with normal imaging, no clinical sign for SOS or GvHD, and no positive biological test for infectious disease. All improved after drug withdrawal. A liver biopsy was performed in ten (7%) patients when liver blood test abnormalities were not explained by clinical, laboratory or imaging results. Diagnosis was established after pathology analysis in nine patients: SOS (n=3), GvHD (n=3), DILI (n=1), HEV infection (n=1) and cholangitis lenta (n=1). One biopsy was not conclusive. Of the eight patients in whom liver biopsy was performed using the transjugular route, six had a hepatic venous pressure gradient (HVPG) <10 mmHg and did not have a SOS, while the two patients with an HVPG >10 mmHg had a confirmed SOS in biopsy.

**Elastography baseline values are not associated with the occurrence of liver involvements after allogeneic hematopoietic stem cell transplantation**

Median basal values before transplantation were 5.4 kPa (interquartile range [IQR], 4.1-6.8) for transient elastography and 6.0 kPa (IQR, 4.9-7.8) for 2D-SWE (Figure 3A). Mean elastography measures from the three radiologists were not significantly different, suggesting that they were not dependent of the operator ($P=0.39$; Online Supplementary Figure S1). TE measure was obtained in 413 out of 432 procedures (two failures at baseline, five at day+7 and 12 at day+14). TE measures did not reach qual-
developed liver involvements after transplantation and those who did not, for both TE (6.4 kPa vs. 6.0 kPa, respectively, P=0.71) and 2D-SWE (7.0 kPa vs. 6.8 kPa, respectively, P=0.71) (Figure SB).

**Sinusoidal obstruction syndrome diagnosis is associated with an increase in two-dimensional real-time shear wave elastography**

In order to determine if repeated measures of elastography could improve or precede clinical diagnosis of liver disease, additional measures of TE and 2D-SWE were systematically performed at day+7 and day+14 after allo-HSCT. Eight (5%) patients were not evaluable at all time points. Three patients had missing exam at day+7 related to intensive care unit (ICU) transfer (n=1) for acute renal failure (without hepatic disorder) or to missing imaging (without hepatic disorder) (n=2). Five patients had missing examination at day+14 related to ICU transfer (n=5) with early death before day+14 (one SOS, one GvHD/congestive heart failure and one with invasive fungal infection) or because of missing data (n=2) for patients with hepatic GvHD (one confirmed by hepatic biopsy and the others with gut GvHD confirmed with pathologic gut biopsy and improvement after immunosuppressive treatment). Liver stiffness value measured using TE or 2D-SWE was neither different at day+7 nor at day+14 between patients developing or not hepatic complications after allo-HSCT (Figure SC and D). Similarly, change in liver stiffness value.
measured using TE or 2D-SWE between baseline and day+7 or day+14 was not associated with hepatic complications after allo-HSCT. We then compared the liver stiffness change from a baseline value to day+7 or day+14 according to each type of liver disease and observed a significant increase only in patients who developed SOS using 2D-SWE at day+14 (+4.15 kPa vs. -0.5 kPa in patients without liver involvement, $P=0.02$) (Figure 3E and F). We also observed a significant increase in patients who developed SOS versus other liver complications using 2D-SWE at day +14 (+4.15 kPa vs. -0.57 kPa, $P=0.018$).

### Two-dimensional real-time shear wave elastography improves positive predictive value of sinusoidal obstruction syndrome diagnosis scores

Fifteen patients (10%) fulfilled EBMT criteria for SOS diagnosis, but only six (4%) had a confirmed SOS diagnosis after retrospective review of medical history (as described in the Methods section) including two late-onset SOS (Online Supplementary Table S1). The other diagnoses were liver GvHD (n=3), sepsis (n=2), cyclosporine cholestasis (n=1), HEV infection (n=1), liver GvHD with congestive heart failure (n=1), and aplastic anemia-related hepatitis with drug-induced liver injury (n=1). According to the Seattle score, twenty-seven (18%) patients had criteria for SOS diagnosis, which was confirmed in only five. Twelve patients (8%) fulfilled SOS criteria according to the Baltimore score, including four confirmed SOS (Figure 4A and B). Median time between allo-HSCT and SOS diagnosis was 14 days (range, 6-22). Two (6%) and four (3.5%) patients with SOS received myeloablative conditioning (MAC) and RIC regimens, respectively. Three of the four SOS patients with RIC regimen ultimately died. Patients were followed until day+100 after allo-HSCT and no other late-onset SOS was detected. Eight patients, including the two late-onset SOS, had a supplementary ultrasound-Doppler. The higher ultrasonography-Doppler score in patient with SOS diagnosis was 2 (n=3), and 5 (n=5, including the 2 late SOS), respectively. The two late-onset SOS patients had an increase of TE value at 36 kPa (vs. 12.4 kPa at baseline) and 72 kPa (vs. 5.4 kPa at baseline). No 2D-SWE was available for these two patients. Patients without SOS (n=6) had a median 2D-SWE value at 4.3 kPa and a median TE value at 6.15 kPa. Among patients with EBMT criteria, most patients with a SOS diagnosis presented an increase of TE and 2D-SWE measures at day+7 or day+14. In four patients, increased 2D-SWE was recorded at a median of 6 days (range, 1-8) before clinical signs of SOS. 2D-SWE increased 1 day after SOS diagnosis in one patient. Among all ultrasonography, Doppler and elastography measures performed at day+7 or day+14, 2D-SWE value at day+14 was the best marker for SOS diagnosis, with a best cutoff value estimated with Youden index at 8.1 kPa ($P=0.004$) and a concordance index calculated at 0.84 (AUROC=0.84 [0.69-0.95]) (Figure 4C). By comparison, TE was less efficient (best cutoff at 8.2 kPa, AUROC=0.78 [0.61-0.91]) (Figure 4D). Ultrasonography and Doppler criteria for SOS were more frequently observed in patients with EBMT criteria but were unable to distinguish between SOS and non-SOS patients (Figure 4E). In order to determine the efficacy of the 2D-SWE measure to improve the predictive value of the current score for SOS diagnosis, we used 3x2 tables as a diagnostic tool to calculate sensibility, specificity and positive predictive value of 2D-SWE combined with clinical criteria. Classical 2x2 tables (after exclusion of non-evaluable measures of 2D-SWE) and 3x2 tables were built to determine the diagnostic value of 2D-SWE measures (Table 3). A worse-scenario approach was calculated, where all non-evaluable measures are classified in the wrong category (absence of SOS diagnosis for patients with SOS and positive SOS diagnosis for patients without SOS). When combined with the EBMT, Baltimore or Seattle scores, a 2D-SWE value higher than 8.1 kPa improved the ability the EBMT score to efficiently diagnose SOS, with a better sensibility (75%) and specificity (99%) for SOS diagnosis (Figure 4F). The addition of 2D-SWE improved the positive predictive value of the Seattle score from 19% to 38%, of the Baltimore score from 33% to 50% and of the EBMT score from 40% to 60%.

### Discussion

After allo-HSCT, liver injury is a major cause of early mortality. Accurate and early diagnosis is the cornerstone for personalized treatment. This study presents a large real-life cohort of 146 consecutive adult patients who
Figure 3. Baseline and evolution of transient elastography and two-dimensional real-time shear wave elastography after allogeneic hematopoietic stem cell transplantation. (A) Individual baseline measure for transient elastography (TE) (5.4 kPa [range, 4.1-6.8]) and two-dimensional real-time shear wave elastography (2D-SWE) (6.0 kPa [range, 4.9-7.8]) with median and interquartile value. Stars represent a patient with fibrosis before transplantation. (B) No significant difference was observed between baseline value (median and interquartile range) of patients with (red) or without (black) liver involvements (unpaired t-test with Holm-Sidak corrections for multiple testing). (C, and D) Sequential evolution of elastography measures compared with baseline with mean and 95% Confidence Interval (CI). No significant difference was observed between patients with (red line) or without (black line) liver involvements for TE (C) and shear wave (D) at day +7 or at day +14 (unpaired t-test with Holm-Sidak correction). For each group, difference between baseline, day +7 and day +14 were not significant (Two-way ANOVA with Dunnet correction for multiple comparison). (E and F) Comparison of elastography delta value (difference between measure and baseline) according to liver diagnosis (box represents median and interquartile values, whiskers are minimum and maximum values) for TE (E) and shear wave (F) (Two-way ANOVA with Dunnet correction multiple comparison). 2D-SWE values were significantly increased in patients with sinusoidal obstruction syndrome (SOS) by comparison with those without liver involvements.
underwent allo-HSCT. In our study, elastography was significantly and specifically increased in patients who developed a SOS. In association with standard SOS scores, 2D-SWE improved the specificity of these scores and their positive predictive value (thus decreasing the risk of over treating patients with expensive drugs with potential side-effects).

In this patient cohort, liver biological abnormalities were frequent after transplantation. Around 39% of patients developed a hyperbilirubinemia, an incidence similar to what has previously been described after allo-HSCT.34 Hyperbilirubinemia has been shown to be predictive of TRM4 and liver injury is associated with a high mortality rate after transplantation.25 We previously reported that clinical and biological features are poorly predictive of liver lesions and that liver biopsy is a useful approach to improve diagnosis of liver involvements in selected patients.3 In this cohort, a low incidence that limits the interpretation of our data on 2D-SWE measures for early diagnosis of SOS. The incidence of an unreliable TE measure (> 10%) could decrease the diagnostic performance in comparison to 2D-SWE. Finally, 2D-SWE was not operator-dependent in our study as in others.35

Many techniques of elastography have been proposed to predict hepatic complications36,37 or to diagnose SOS after allo-HSCT.36,38,39 In a pediatric cohort of 22 patients, sequential measures of TE after allo-HSCT showed increased TE measures 3 to 6 days before clinical signs in five SOS cases (based on Seattle or Baltimore classification).38 In a second pediatric series of 25 patients, sequential 2D-SWE velocity at day+5 and day+14 in five SOS patients was higher than in the control population.39 In the adult setting, one single center study including 78 patients after allo-HSCT evaluated the usefulness of sequential measures of TE for SOS diagnosis.40 TE was significantly increased in four patients with SOS.

Currently, SOS diagnosis relies on clinical criteria that have low specificity. To our knowledge, there is no published evaluation of the sensibility, specificity or predictive value of EBMT, Seattle or Baltimore criteria, using liver biopsy as a gold standard method for SOS diagnosis.4 In our cohort, one third of patients with EBMT criteria had a liver biopsy and the systematic retrospective review of their medical history helped us to better estimate the predictive value of each clinical score.41 For patients who did
Figure 4. Two-dimensional real-time shear wave elastography measures improve sinusoidal obstruction syndrome diagnosis. (A and B) Venn diagram representing the distribution of patients with European Society for Blood and Marrow Transplantation (EBMT), Seattle and/or Baltimore Criteria in this cohort (A) and those with final diagnosis of sinusoidal obstruction syndrome (SOS) in each group (B). (C and D) ROC curves for day+14 two-dimensional real-time shear wave elastography (2D-SWE) and transient elastography (TE) measures, and box plot of individual values (box represents median and interquartile values, whiskers are minimum and maximum values) in patients with (blue) or without SOS (yellow) for shear wave (A) and TE (B) compared with Mann-Whitney U test. (E) Radar plot of individual ultrasound criteria and box plot of ultrasonography scores (calculated by adding one point per criteria) for patients with no EBMT criteria (gray lines) and with EBMT criteria without (orange lines) or with SOS diagnosis (blue lines) compared with Kruskal Wallis test and Dunn’s correction for multiple testing. (F) Representation of sensitivity, specificity, and positive predictive rate (PPR) (color circle, large and red circle are associated with greatest PPR). Addition of 2D-SWE measure with a cutoff value at 8.1 kPa improved the PPR value of the Seattle score from 19% to 38%, of the Baltimore score from 33% to 50% and of the EBMT score from 40% to 60%.
not have a biopsy, the final diagnosis relied on a retrospective analysis of all medical data, of clinical evolution with treatments, and on consensual diagnosis criteria if available. However, the lack of gold standard criteria for most of diagnoses, such as GvHD, SOS or DILI, could bias conclusions as there is currently no reliable tool to avoid misdiagnoses. An unreliable TE or 2D-SWE value is classically estimated to be included between 5% and 10%, a range very similar to what was observed in our cohort. Using an intention to diagnose approach helped us to estimate the suitability of the liver stiffness measurement in addition to clinical criteria for SOS diagnosis. Our results confirm the interest of liver stiffness measures and strongly suggest that 2D-SWE significantly improved the positive predictive value of clinical scores for SOS diagnosis after allo-HSCT. Increasing the positive predictive value of current criteria for SOS is critical to avoid useless and potentially toxic treatment in patient with comorbidities.

Thus, 2D-SWE appears as a promising, non-invasive, quantitative, safe, and reproducible technique allowing an early and accurate diagnosis of SOS. 2D-SWE measures combined with classical scores such as Baltimore or EBMT criteria could significantly help discriminating SOS from other post-transplantation early liver injury. Prospective multi-center trials would be necessary to confirm these results and to broadly evaluate the impact of liver stiffness measures on management and treatment of liver involvement after allo-HSCT.

Disclosures
No conflicts of interest to disclose.

Contributions
PED and DM collected, analyzed data and performed statistics analysis; PB, AMZ, and MDB performed ultrasonography, Doppler and elastography measurements; PER, SP, ASG, MR, RPL, AP, FSE, AX, PHP and DV provided data and commented manuscript; PED, GS and DM wrote the manuscript; GS and DM conceived the project and supervised the work; all authors approved the manuscript.

References


