

Training cohort: clinicopathologic characteristics and follow-up

The training cohort was prospectively accrued within a timeframe 18 months and consisted of 57 patients that underwent ERCP for a biliary duct stricture (detailed information in Supplementary Table 1). Patients ranged in age from 26 to 96 years (mean, 66.7 years; median, 69.0 years) and were predominantly male (30 of 57, 53%). A history of PSC was noted for 12 (21%) patients. There were multiple indications for ERCP and included symptomatic presentation (54 of 57, 95%), which was characterized by obstructive jaundice, abdominal pain, pruritus, weight loss and/or fever; abnormal pancreatobiliary radiographic imaging (n = 39, 68%); elevated bilirubin and/or abnormal liver function tests (n = 34, 60%); and stent exchange (n = 18, 32%). Twenty-two (39%) patients had a previous ERCP. Serum CA19-9 was measured at the time of ERCP for 56 (98%) cases and ranged between 0.8 and 34,793 Units per milliliter (mL) (mean, 1,831 U/mL; median, 38.6 U/mL).

For each of the 57 patients, a single gastroenterologist (A.S.) performed the ERCP. Biliary duct strictures varied in location and consisted of 29 (51%) distal, 12 (21%) hilar and 16 (28%) intrahepatic strictures. On the basis of the Bismuth-Corlett classification system, hilar strictures were further subdivided into 4 Bismuth I, 2 Bismuth II, 4 Bismuth IIIa, 1 Bismuth IIIb and 1 Bismuth IV strictures. Among the 57 patients, specimens submitted for pathologic evaluation included biliary brushings for 19 (33%) patients, biliary biopsies for 25 (44%) patients and both biliary brushings and biopsies for 13 (23%) patients. Among the biliary biopsies, 13 of 38 (34%) specimens were obtained by cholangioscopy, while the remaining specimens were acquired through cholangiography. A paired but separate brushing and/or biopsy specimen was also submitted for BiliSeq testing. However, among the 13 patients that had both brushing and biopsy specimens for pathologic evaluation, only 2 patients had paired brushing and cholangiographic-

guided biopsy specimens obtained for BiliSeq testing. For the remaining 11 patients, a corresponding brushing specimen, but not biopsy specimen, was submitted for BiliSeq testing. In total, 59 bile duct specimens from 57 patients were evaluated by BiliSeq.

A repeat ERCP was performed for 5 of 57 (9%) patients (summarized in Supplementary Table 2) with 1 patient undergoing repeat ERCP three times. Among these 7 repeat procedures, either biliary cholangiographic-guided biopsies ($n = 5$), brushings ($n = 1$) or both cholangiographic-guided biopsies and brushings ($n = 1$) were submitted for pathologic evaluation on repeat ERCP. For the 1 patient that underwent repetitive ERCP, only biliary biopsies were obtained for pathologic evaluation. For the patients that only had biliary biopsies or brushings, a parallel specimen was obtained for BiliSeq testing; while the patient that had both a biopsy and brushings for pathologic evaluation, only a brushing specimen was submitted for BiliSeq testing. Follow-up was available for all 57 (100%) patients and consisted of diagnostic pathology for 34 (61%) patients and ≥ 12 months of clinicoradiographic follow-up for 23 (41%) patients. Twenty-six of 57 (46%) patients died during the study follow-up period. The final diagnoses included 12 distal cholangiocarcinomas, 12 perihilar cholangiocarcinomas, 5 intrahepatic cholangiocarcinomas, 3 pancreatic ductal adenocarcinomas, 3 gallbladder adenocarcinomas and 22 benign cholangiopathies.

Validation cohort: clinicopathologic characteristics and follow-up

The validation phase of this study consisted of 195 patients with a biliary duct stricture that were prospectively evaluated over the course of 22 months by several gastroenterologists at the UPMC Digestive Disorders Center (detailed information is provided in Supplementary Table 1). None of the patients within the validation cohort were included within the training cohort. Age at

presentation ranged between 15 and 96 years (mean, 66.4 years; median, 70.0 years) and the patients were predominantly male (110 of 195, 56%). Twenty-nine (15%) patients had a documented history of PSC. Indications for ERCP included symptomatic presentation (n = 169, 87%), abnormal pancreatobiliary radiographic imaging (n = 117, 63%), elevated bilirubin and/or abnormal liver function tests (n = 121, 62%) and stent exchange (n = 28, 14%). An ERCP was previously performed on 39 (20%) patients. At the time of ERCP, serum CA19-9 studies were obtained for 162 (83%) patients and ranged from 0.8 and 20,460 U/mL (mean, 1,288.5 U/mL; median, 99.9 U/mL).

The biliary duct strictures were located within the following anatomic sites: 137 (70%) distal, 35 (18%) hilar and 23 (12%) intrahepatic. Among the hilar strictures, 3 patients had a Bismuth I stricture, 10 patients had Bismuth II stricture, 9 patients had a Bismuth IIIa stricture, 3 patients had a Bismuth IIIb stricture and 10 patients had a Bismuth IV stricture. Specimens submitted for pathologic evaluation included biliary brushings for 57 (29%) patients, biliary biopsies for 64 (33%) patients and both biliary brushings and biopsies for 74 (38%) patients. Among the biliary biopsies, 29 of 138 (21%) specimens were obtained by cholangioscopy, while the remaining specimens were acquired through cholangiography. A paired but separate brushing and/or biopsy specimen was also submitted for BiliSeq testing. For the 74 patients that had both brushing and biopsy specimens for pathologic evaluation, 41 (55%) also had separate brushing and biopsy specimens submitted for BiliSeq testing. Among the remaining 33 patients, corresponding brushings or biopsies were obtained for 27 patients and 6 patients, respectively, for BiliSeq testing. In total, 236 bile duct specimens from 195 patients were submitted for BiliSeq testing.

Twenty-one (11%) patients had a repeat ERCP with 3 patients undergoing repeat ERCP 2 times and 2 patients undergoing repeat ERCP 3 times. Among these 28 repeat procedures, specimens submitted for pathologic evaluation included 6 biliary biopsies, 3 brushings and 19 pairs of both biopsies and brushings. Among the biliary biopsies, 4 of 25 (16%) specimens were obtained by cholangioscopy, while the remaining specimens were acquired through cholangiography. For the procedures where either a biliary biopsy or brushing was obtained, a parallel specimen was submitted for BiliSeq testing. Sixteen of the 19 pairs of biopsies and brushings submitted for pathologic evaluation also had corresponding biopsies and brushings obtained for BiliSeq testing. A brushing specimen was submitted for BiliSeq testing for the remaining 3 procedures. Follow-up was available for 163 of 195 (84%) patients and consisted of diagnostic pathology for 111 (68%) patients and ≥ 12 months of clinicoradiographic follow-up for 52 (32%) patients. Sixty of 163 (37%) patients died during the study follow-up period. The final diagnoses included 35 distal cholangiocarcinomas, 14 perihilar cholangiocarcinomas, 17 intrahepatic cholangiocarcinomas, 30 pancreatic ductal adenocarcinomas, 8 gallbladder adenocarcinomas, 7 ampullary adenocarcinomas, 1 mixed hepatocellular carcinoma-cholangiocarcinoma, 1 metastatic colonic adenocarcinoma, 2 high-grade dysplasias of the bile duct and 48 benign cholangiopathies.

BiliSeq testing results and clinicopathologic correlation

The DNA concentrations from submitted ERCP-obtained bile duct specimens ranged from 0.08 to 35.71 ng/uL (mean, 4.45 ng/uL; median, 2.18 ng/uL) and all 346 specimens (biliary brushings, n = 184; biliary biopsies, n = 162) were sufficient for BiliSeq testing. Concentrations of DNA did not vary significantly between bile duct specimens with biliary brushings ranging

from 0.08 to 27.45 ng/uL (mean, 3.17 ng/uL; median, 1.61 ng/uL) and biliary biopsies ranging from 0.53 to 35.71 ng/uL (mean, 6.42 ng/uL; median, 4.22 ng/uL). Overall genomic alterations were detected in 69 of 184 (38%) brushings and 81 of 162 (50%) biopsies and consisted of 113 unique alterations in 18 genes. The most frequent genomic alteration involved *KRAS* with a prevalence of 24% (82 of 346). Additional genomic alterations in decreasing order occurred in *TP53* (n = 74, 21%), *CDKN2A* (n = 22, 6%), *SMAD4* (n = 18, 5%), *PIK3CA* (n = 14, 4%), *GNAS* (n = 11, 3%), *CTNNB1* (n = 7, 2%), *ERBB2* (n = 7, 2%), *FGFR2* (n = 5, 1%), *BRAF* (n = 4, 1%), *MET* (n = 4, 1%), *ALK* (n = 3, 1%), *ATM* (n = 3, 1%), *FGFR3* (n = 2, 1%), *IDH2* (n = 2, 1%) and *IDH1* (n = 1, 1%). Univariate analysis revealed the presence of a genomic alteration was associated with an increased mean age (70.5 years versus 64.9 years, $p = 0.001$), a serum CA 19-9 of ≥ 44 U/mL ($p < 0.001$), and a concurrent diagnosis of suspicious for adenocarcinoma on brushings and/or at least high-grade dysplasia on biopsies ($p < 0.001$). There was no statistically significant correlation between the presence of a genomic alteration and patient gender ($p = 0.721$), location of the bile duct stricture ($p = 0.842$) and a history of PSC ($p = 0.058$). Moreover, no differences were identified among hilar strictures based on the Bismuth-Corlett classification system ($p = 0.923$).