



Integrating next-generation sequencing to endoscopic retrograde cholangiopancreatography (ERCP)-obtained biliary specimens improves the detection and management of patients with malignant bile duct strictures

Aatur D Singhi ,¹ Marina N Nikiforova,¹ Jennifer Chennat,² Georgios I Papachristou,² Asif Khalid,² Mordechai Rabinovitz,² Rohit Das,³ Savreet Sarkaria,³ M Samir Ayasso,³ Abigail I Wald,¹ Sara E Monaco,¹ Michael Nalesnik,⁴ N Paul Otori,¹ David Geller,⁵ Allan Tsung,⁵ Amer H Zureikat,⁵ Herbert Zeh,⁶ J Wallis Marsh,⁷ Melissa Hogg,⁸ Kenneth Lee,⁹ David L Bartlett,⁵ James F Pingpank,⁵ Abhinav Humar,¹⁰ Nathan Bahary,¹¹ Anil K Dasyam,¹² Randall Brand,² Kenneth E Fasanella,² Kevin McGrath,² Adam Slivka²

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2018-317817>).

For numbered affiliations see end of article.

Correspondence to

Dr Aatur D Singhi, Department of Pathology, University of Pittsburgh Medical Center Health System, Pittsburgh, PA 15213, USA; singhiad@upmc.edu

Received 29 October 2018

Revised 25 March 2019

Accepted 27 March 2019

Published Online First

10 April 2019



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Singhi AD, Nikiforova MN, Chennat J, et al. *Gut* 2020;**69**:52–61.

ABSTRACT

Objective Despite improvements in imaging, serum CA19-9 and pathological evaluation, differentiating between benign and malignant bile duct strictures remains a diagnostic conundrum. Recent developments in next-generation sequencing (NGS) have opened new opportunities for early detection and management of cancers but, to date, have not been rigorously applied to biliary specimens.

Design We prospectively evaluated a 28-gene NGS panel (BiliSeq) using endoscopic retrograde cholangiopancreatography-obtained biliary specimens from patients with bile duct strictures. The diagnostic performance of serum CA19-9, pathological evaluation and BiliSeq was assessed on 252 patients (57 trainings and 195 validations) with 346 biliary specimens.

Results The sensitivity and specificity of BiliSeq for malignant strictures was 73% and 100%, respectively. In comparison, an elevated serum CA19-9 and pathological evaluation had sensitivities of 76% and 48%, and specificities of 69% and 99%, respectively. The combination of BiliSeq and pathological evaluation increased the sensitivity to 83% and maintained a specificity of 99%. BiliSeq improved the sensitivity of pathological evaluation for malignancy from 35% to 77% for biliary brushings and from 52% to 83% for biliary biopsies. Among patients with primary sclerosing cholangitis (PSC), BiliSeq had an 83% sensitivity as compared with pathological evaluation with an 8% sensitivity. Therapeutically relevant genomic alterations were identified in 20 (8%) patients. Two patients with *ERBB2*-amplified cholangiocarcinoma received a trastuzumab-based regimen and had measurable clinoradiographic response.

Conclusions The combination of BiliSeq and pathological evaluation of biliary specimens increased the detection of malignant strictures, particularly in patients with PSC. Additionally, BiliSeq identified

Significance of this study

What is already known on this subject?

- Despite a multidisciplinary approach, distinguishing between benign and malignant bile duct strictures can be challenging.
- Whole exome and whole genome sequencing studies have defined the genetic landscape of neoplasms arising and secondarily involving the bile duct system.
- A test that could detect key genomic alterations in malignant neoplasms involving the bile duct system may be of diagnostic utility to endoscopic retrograde cholangiopancreatography (ERCP)-obtained biliary specimens.

alterations that may stratify patients for specific anticancer therapies.

INTRODUCTION

Strictures involving the bile duct represent a heterogeneous group of benign and malignant conditions.¹ Benign bile duct strictures may be due to primary sclerosing cholangitis (PSC), IgG4-related sclerosing cholangitis, iatrogenic injury, infection and other less common aetiologies. Malignant causes include carcinomas involving the pancreatobiliary ducts, ampulla of Vater, liver and, rarely, metastatic cancers. However, differentiating between benign and malignant bile duct strictures can be challenging and relies on a multidisciplinary approach of clinical examination, biochemical testing (eg, serum CA19-9), radiographic imaging, endoscopic procedures and pathological evaluation with

Significance of this study

What are the new findings?

- ▶ We have developed and validated a highly sensitive, targeted next-generation sequencing (NGS) assay (BiliSeq) within a Clinical Laboratory Improvement Amendments-certified and College of American Pathologists-accredited laboratory for 28 genes that are commonly mutated, amplified and/or deleted in malignant neoplasms involving the bile duct system.
- ▶ A large prospective analysis of ERCP-obtained biliary specimens revealed BiliSeq improved the sensitivity of pathological evaluation for both biliary brushing and biliary biopsy specimens.
- ▶ Among patients with primary sclerosing cholangitis, BiliSeq was superior to serum CA19-9 and pathological evaluation in detecting at least high-grade biliary dysplasia.
- ▶ BiliSeq identified potentially actionable genomic alterations with demonstrable chemotherapeutic response for a subset of patients.

How might it impact on clinical practice in the foreseeable future?

- ▶ These results highlight the diagnostic applicability of NGS-based assays to ERCP-obtained biliary specimens in early detection and management of patients with malignant bile duct strictures.

ancillary studies.² This distinction is more arduous considering the aetiology of some benign strictures can be a predisposing factor for malignancy. For example, PSC is associated with a 400-fold increased risk of cholangiocarcinoma as compared with the general population.

Endoscopic retrograde cholangiopancreatography (ERCP) plays a key role in the evaluation of a bile duct stricture by defining the morphometric imaging features and extent of disease involvement.³ However, there are no pathognomonic features seen during fluoroscopy that can differentiate benign from malignant. Furthermore, direct cholangioscopic grading is inaccurate and suffers from poor interobserver agreement.^{4,5} During ERCP, bile duct brushings and forceps biopsies are the methods of choice for pathological confirmation; however, the sensitivity of detecting malignancy can range between 8% and 67%.^{4,6–11} To improve the detection of malignant strictures, adjunctive techniques were developed and include digital image analysis, *KRAS* mutational testing and fluorescence in situ hybridisation.^{12–18} However, the sensitivities of these assays range between 14% and 60%. Considering the low sensitivity of pathological evaluation and available ancillary studies, it is not surprising that patients often undergo multiple ERCP procedures for diagnostic purposes. As a result, therapeutic decisions for malignant strictures are often delayed by several weeks to months, which places patients at risk for disease progression. Conversely, a misdiagnosis of a benign bile duct stricture can result in unnecessary surgical resection. It is reported that up to 15% of patients undergoing surgery for a suspected malignant stricture have benign disease.^{19,20} This is alarming given these surgical resections are associated with a relatively, high morbidity and mortality.^{21–24} Thus, more innovative preoperative approaches are needed to improve the stratification of patients with benign and malignant strictures.

Over the past several years, dramatic progress has been made in understanding the genomic landscape of neoplasms arising

or secondarily involving the bile duct system. Whole exome and whole genome sequencing studies have reported recurrent genomic alterations in a relatively small number of oncogenes and tumour suppressor genes, such as *KRAS*, *TP53*, *CDKN2A* and *SMAD4*.^{25–34} Moreover, a subset of these genomic alterations, such as *ATM* and *ERBB2*, may confer susceptibility to specific anticancer therapies. In parallel, the introduction of novel molecular diagnostics has opened new opportunities for the study of preoperative specimens.^{35,36} Bile duct specimens are often characterised by small quantities of diagnostic material and composed of heterogeneous cell populations, which may obscure or mimic a malignant process. It is therefore crucial for a molecular assay to be highly sensitive to detect small proportions of mutated cells within these specimens. Next-generation sequencing (NGS) combines a high analytical sensitivity with multigene analysis and thus represents an attractive option for the assessment of bile duct strictures.

In this study, we developed a highly sensitive, targeted NGS assay, known as BiliSeq, for 28 genes that are commonly mutated, amplified and/or deleted in malignant neoplasms involving the bile duct system.^{25–34,37} This test was performed within a Clinical Laboratory Improvement Amendments (CLIA)-certified and College of American Pathologists (CAP)-accredited clinical laboratory using ERCP-obtained bile duct brushings and/or biopsies. Rather than extracting DNA from cytological smears/slides, alcohol fixative (eg, CytoLyte) or formalin-fixed paraffin-embedded (FFPE) tissue, which may decrease overall yields and quality, a dedicated bile duct brushing and/or biopsy was submitted directly for BiliSeq testing after routine specimens were obtained for pathological evaluation. Our objectives were to prospectively evaluate BiliSeq testing on a large cohort of patients to: (1) determine its accuracy in the detection of malignant strictures, (2) compare its performance as an adjunct to other diagnostic modalities and (3) evaluate the ramifications on patient management when genomic alterations are detected in bile duct specimens.

MATERIALS AND METHODS**Patient population**

A training cohort of biliary specimens from 57 patients was prospectively collected by standard ERCP at the University of Pittsburgh Medical Center (UPMC) from patients with a bile duct stricture between September 2014 and April 2016 and funded by the University of Pittsburgh Department of Pathology. On completion of the training cohort, a validation cohort of 195 patients was prospectively accrued between September 2016 and August 2018 and partially funded by the Institute for Precision Medicine at the University of Pittsburgh. In total, 163 biliary brushings and 172 biliary biopsies were collected for pathological evaluation, and 160 brushings and 135 biopsies were collected for BiliSeq testing. Bile duct specimens were submitted to the UPMC Anatomic Pathology laboratory for routine cytopathological and pathological evaluation and the UPMC Molecular & Genomic Pathology (MGP) Laboratory for BiliSeq testing within 24–48 hours of the ERCP procedure. Medical records were reviewed to document patient demographics, clinical presentation, ERCP findings, concurrent endoscopic ultrasound (EUS) and EUS-fine needle aspiration (FNA), serum CA19-9 levels and pathological diagnoses of corresponding brushings and/or biopsies. A separate analysis was performed for repeat ERCP procedures that yielded 24 biliary brushings and 29 biliary biopsies for pathological evaluation and 24 brushings and 27 biopsies for

BiliSeq testing. The aetiology of the stricture was based on diagnostic pathology (n=145) from surgical resected specimens or biopsy specimens (either autopsy, open or laparoscopic surgical resection/biopsy, percutaneous needle or ERCP-obtained biliary biopsy) and/or clinical criteria (n=75) after ≥ 12 months of follow-up. As the reported specificity of biliary brushings is $< 100\%$, the findings from pathological evaluation of brushing specimens were not considered as diagnostic pathology. Furthermore, a negative ERCP-obtained biopsy was not considered supportive of a benign aetiology. A stricture was designated as a benign cholangiopathy if a repeat imaging or ERCP at ≥ 12 months follow-up documented the resolution or stability of prior ductal abnormalities. In contrast, a clinical diagnosis of malignancy was defined by the combination of a mass on radiographic imaging in the absence of acute cholangiopathy, and clinical or radiographic progression after ≥ 12 months of follow-up, or death clinically and radiographically determined to be due to a malignancy involving the bile duct. For surgical resection and biopsy specimens, diagnoses were rendered on the basis of standard histomorphological criteria.³⁸ Diagnostic and clinical follow-up information including the method of follow-up for each patient is described in detail within online supplementary table 1.

Specimen collection

Samples were obtained during the ERCP procedure and included cytological brushings, forceps biopsies or both. For cytological brushes, samples were collected with multiple to-and-fro motions at the site of the stricture. The brush used

for cytopathological examination and was placed into 15 mL of ThinPrep CytoLyt solution (Hologic, Marlborough, Massachusetts, USA) and transferred for specimen processing within 24 hours. A separate brush was used for BiliSeq testing and placed into a collection vial containing a detergent-based DNA lysis buffer. The collection vial was then transported to the UPMC MGP laboratory within 24 hours and processed accordingly. A similar approach was used for forceps biopsies with separate biopsies submitted for pathological examination and BiliSeq testing. No minimum specimen cellularity was required for testing. All specimens whether a brushing or biopsy were processed the same.

BiliSeq testing

NGS was performed prospectively as part of clinical care and within a 10-day turnaround in the UPMC MGP laboratory, which is both CLIA-certified and CAP-accredited. Genomic DNA was isolated similarly for both brushing and biopsy specimens using the MagNA Pure LC Total Nucleic Acid Isolation Kit (Roche, Indianapolis, Indiana, USA) on Compact MagNA Pure (Roche). In cases where both a brushing and biopsy specimen was submitted, the DNA from both specimens was combined for subsequent analysis. Extracted DNA was quantified on the Qubit 2.0 Fluorometer using the dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Amplification-based targeted NGS was performed with primers for genomic regions of interest (https://mgp.upmc.com/Applications/mgp/Content/OncoSeq_Hotspots.pdf) within the following genes: *AKT1*, *ALK*, *ATM*, *BRAF*,

Table 1 Clinicopathological characteristics of training and validation cohorts based on the absence (negative) versus presence (positive) of genomic alterations

Patient or bile duct characteristics	Total, n=252	Training cohort		P value	Validation cohort		P value
		Negative (n=35, 61%)	Positive (n=22, 39%)		Negative (n=104, 53%)	Positive (n=91, 47%)	
Gender							
Female	112	17 (63)	10 (37)	1.000	47 (55)	38 (45)	0.666
Male	140	18 (60)	12 (40)		57 (52)	53 (48)	
Mean age (range), years	67.2 (15–96)	65.5 (26–96)	68.7 (51–92)	0.469	64.3 (15–89)	71.0 (33–96)	< 0.001
History of PSC	41	8 (67)	4 (33)	0.750	23 (79)	6 (21)	0.002
CA 19–9 ≥ 44 U/mL (n=218)*	134	10 (47)	16 (53)	0.002	47 (44)	61 (56)	0.013
CA 19–9 > 129 U/mL (n=218)*	96	6 (32)	13 (68)	0.002	28 (36)	49 (64)	0.001
Location of stricture							
Intrahepatic	38	9 (60)	6 (40)	1.000	13 (57)	10 (43)	0.219
Hilar	47	7 (58)	5 (42)		14 (40)	21 (60)	
Distal	167	19 (63)	11 (37)		77 (56)	60 (44)	
Malignant biliary brushing/biopsy†	74	5 (31)	11 (69)	0.006	12 (21)	46 (79)	< 0.001
Clinical and diagnostic pathology follow-up	n=220	n=35	n=22		n=76	n=87	
Intrahepatic cholangiocarcinoma	22	0 (0)	5 (100)	< 0.001 †	5 (29)	12 (17)	< 0.001 †
Perihilar cholangiocarcinoma	26	5 (42)	7 (58)		3 (21)	11 (79)	
Distal cholangiocarcinoma	47	6 (50)	6 (50)		11 (31)	24 (69)	
Mixed HCC-cholangiocarcinoma	1	–	–		0 (0)	1 (100)	
Pancreatic ductal adenocarcinoma	33	0 (0)	3 (100)		6 (20)	24 (80)	
Gallbladder adenocarcinoma	11	2 (67)	1 (33)		3 (37)	5 (63)	
Ampullary adenocarcinoma	7	–	–		0 (0)	7 (100)	
Metastatic colonic adenocarcinoma	1	–	–		0 (0)	1 (100)	
High-grade biliary dysplasia	2	–	–		0 (0)	2 (100)	
Benign cholangiopathy	70	22 (100)	0 (0)		48 (100)	0 (0)	

*Serum CA19-9 was measured for 218 of 252 (87%) patients.

†Malignant biliary brushing/biopsy is defined as at least high-grade dysplasia/suspicious for adenocarcinoma. CA19-9, carbohydrate antigen 19-9; HCC, hepatocellular carcinoma; PSC, primary sclerosing cholangitis.

CDKN2A, CTNNB1, EGFR, ERBB2, ERBB4, FGFR1, FGFR2, FGFR3, GNA11, GNAQ, GNAS, HRAS, IDH1, IDH2, KIT, KRAS, MET, NRAS, PDGFRA, PIK3CA, PTEN, SMAD4, TP53 and *VHL* as previously described.³⁹ Amplicons were barcoded, purified and ligated with specific adapters. DNA quantity and quality check were performed using the 2200 TapeStation (Agilent Technologies, Santa Clara, California, USA). The Ion One Touch 2 and One Touch ES were used to prepare and enrich templates and enable testing via Ion Sphere Particles on a semiconductor chip. Massive parallel sequencing was carried out on an Ion Proton System according to the manufacturer's instructions (Thermo Fisher Scientific) and analysed with the Torrent Suite Software V.3.4.2 for point mutations, small insertions/deletions and copy number alterations. Each variant was prioritised according to the 2017 Association for Molecular Pathology (AMP)/American Society of Clinical Oncology (ASCO)/College of American Pathologists (CAP) joint consensus guidelines for interpretation of sequence variants in cancer using a tier-based system.⁴⁰ Tier I, Tier II and Tier III variants were reported; however, only Tier I and Tier II variants were used for subsequent analysis. The presence of a genomic alteration that qualified as a Tier I or Tier II variant was considered to be a 'positive' BiliSeq result, while a 'negative' BiliSeq result was determined to be the absence of a Tier I or Tier II variant. The limit of detection of the assay was at 3% mutant allele frequency (AF). The minimum depth of coverage for testing was 500×. For each mutation identified, a AF was calculated based on the number of reads of the mutant allele versus the wild-type allele and reported as a percentage.³⁵ Copy number variation analysis was performed as previously described.⁴¹ A gene amplification was defined by the presence of ≥6 copies of a variant as previously described and validated using fluorescence in situ hybridisation (FISH) analysis.^{41 42}

Classification of targetable genomic alterations

Clinically relevant genomic alterations were defined as alterations that are targetable by anticancer drugs currently available on the market or in registered clinical trials. Genomic alterations that are potentially targetable with reported kinase inhibitors (in the setting of wild-type *KRAS*, *NRAS* and *HRAS* genomic alterations) include: *ALK*, *BRAF*, *EGFR*, *ERBB2*, *FGFR1*, *FGFR2*, *FGFR3* and *MET*.^{43–46} Genomic alterations within *ATM* may be targeted with platinum-based regimens and poly-(ADP-ribose) polymerase inhibition.⁴⁷ Also, pharmacological inhibitors highly specific to individual *IDH1* and *IDH2* mutations have been developed.^{48 49}

Statistical analysis

The sample size estimate/power calculation for the training cohort used within this study was based on a 60% prevalence of malignancy among bile duct strictures at our institution and corresponding 40% sensitivity for detecting malignancy using pathological evaluation of ERCP-obtained specimens. Therefore, a sample size of at least 50 patients would yield a >80% power with an alpha of 0.05. Statistical analyses to assess differences in mutational status were compared using Fisher's exact test for dichotomous variables. For the training cohort, receiver operating characteristic (ROC) curves were generated to compare biomarker performances and to establish cut-off levels, using the Youden index. Final cut-offs were selected to maximise the area under the curve (AUC) as calculated by the trapezoidal method. However, biological rationale and the simplicity of the model were also considered. Sensitivity and specificity were calculated using standard 2×2 contingency tables for cases with confirmed diagnostic pathology. Unless otherwise stated, repeat testing was not used to calculate sensitivity and specificity for individual

Table 2 Sensitivities and specificities of serum CA19-9, biliary brushings/biopsies and BiliSeq for at least high-grade biliary dysplasia

Parameter	Sensitivity, n (95% CI)	Specificity, n (95% CI)	PPV, n (95% CI)	NPV, n (95% CI)	AUC, n (95% CI)
Training cohort (n=57)					
Elevated CA19-9 (n=56)*	62 (0.44 to 0.77)	77 (0.54 to 0.91)			
Pathological evaluation of biliary brushings/biopsies	46 (0.29 to 0.63)	100 (0.82 to 1.00)			
BiliSeq testing	63 (0.45 to 0.78)	100 (0.82 to 1.00)			
Elevated CA19-9* and BiliSeq testing	79 (0.62 to 0.91)	77 (0.54 to 0.91)			
Pathological evaluation and BiliSeq testing	77 (0.60 to 0.89)	100 (0.82 to 1.00)			
Elevated CA19-9*, pathological evaluation and BiliSeq testing	88 (0.72 to 0.96)	77 (0.54 to 0.91)			
Validation cohort (n=163)					
Elevated CA19-9 (n=144)*	81 (0.72 to 0.88)	65 (0.48 to 0.79)			
Pathological evaluation of biliary brushings/biopsies	49 (0.39 to 0.58)	98 (0.88 to 1.00)			
BiliSeq testing	76 (0.67 to 0.83)	100 (0.91 to 1.00)			
Elevated CA19-9* and BiliSeq testing	97 (0.91 to 0.99)	65 (0.48 to 0.79)			
Pathological evaluation and BiliSeq testing	84 (0.76 to 0.90)	98 (0.88 to 1.00)			
Elevated CA19-9*, pathological evaluation and BiliSeq testing	97 (0.91 to 0.99)	63 (0.46 to 0.77)			
Overall cohort (n=220)					
Elevated CA19-9 (n=200)*	76 (0.68 to 0.83)	69 (0.56 to 0.80)	85 (0.77 to 0.90)	57 (0.45 to 0.68)	0.73 (0.65 to 0.81)
Pathological evaluation of biliary brushings/biopsies	48 (0.40 to 0.56)	99 (0.91 to 1.00)	99 (0.92 to 1.00)	47 (0.39 to 0.55)	0.73 (0.67 to 0.80)
BiliSeq testing	73 (0.65 to 0.80)	100 (0.94 to 1.00)	100 (0.96 to 1.00)	63 (0.53 to 0.72)	0.86 (0.82 to 0.91)
Elevated CA19-9* and BiliSeq testing	93 (0.87 to 0.96)	69 (0.56 to 0.80)	87 (0.80 to 0.92)	81 (0.68 to 0.90)	0.81 (0.74 to 0.88)
Pathological evaluation and BiliSeq testing	83% (0.75 to 0.88)	99 (0.91 to 1.00)	99 (0.95 to 1.00)	73 (0.62 to 0.81)	0.91 (0.87 to 0.95)
Elevated CA19-9*, pathological evaluation and BiliSeq testing	95 (0.89 to 0.98)	68 (0.55 to 0.79)	87 (0.80 to 0.92)	86 (0.72 to 0.94)	0.81 (0.74 to 0.89)

*On the basis of cases where serum CA19-9 was performed concurrently where elevated CA19-9 was defined as ≥44 U/mL. AUC, area under the curve; CA19-9, carbohydrate antigen 19-9; NPV, negative predictive value; PPV, positive predictive value.

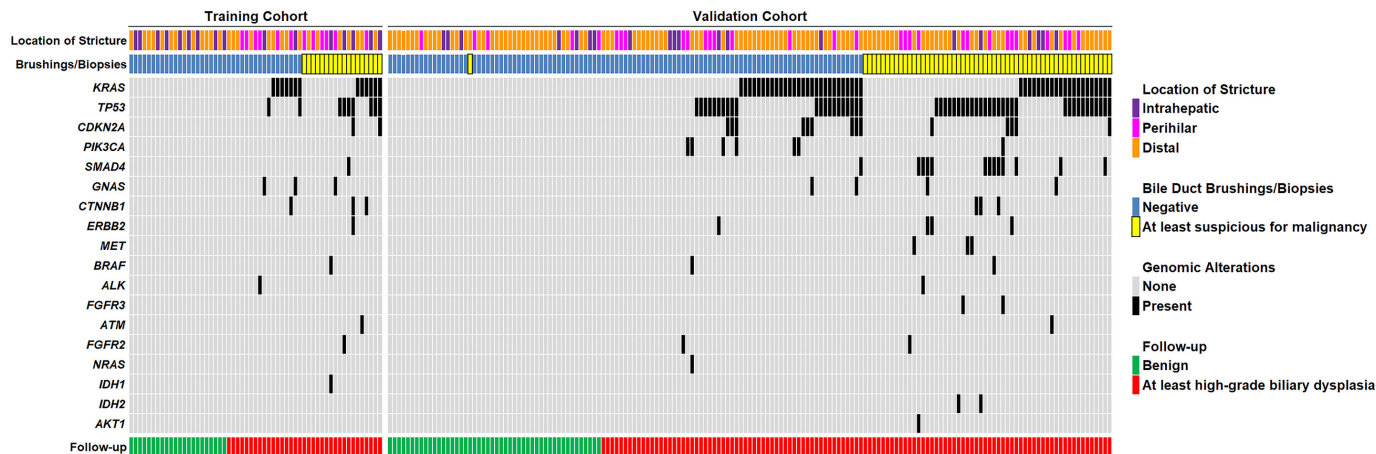


Figure 1 Correlative findings of both training and validation cohorts for pathological examination of biliary brushings/biopsies and individual genomic alterations from BiliSeq testing with corresponding patient follow-up. Genomic alterations identified through BiliSeq testing were associated with a higher sensitivity (74% vs 50%) and higher specificity (100% vs 99%) for at least high-grade biliary dysplasia involving the bile duct system. In decreasing order, the most prevalent genomic alterations included *KRAS*, *TP53*, *CDKN2A*, *SMAD4*, *PIK3CA*, *GNAS* and others.

biomarkers. All statistical analyses were performed using the SPSS Statistical software, V.25 (IBM, Armonk, New York, USA) and statistical significance was defined as a p value of <0.05.

RESULTS

Performance of serum CA19-9, pathological evaluation and BiliSeq testing within training and validation cohorts

A training cohort of 57 patients that underwent ERCP for a bile duct stricture were prospectively accrued for biomarker analysis and summarised in [table 1](#), online supplementary data and table 1. Serum CA19-9 studies, bile duct brushings and/or biopsies for pathological evaluation and BiliSeq testing were performed for 56 of 57 (98%) patients with only one patient that did not have a CA19-9 measurement at the time of clinical presentation. Considering the lack of a consensus CA19-9 cut-off as a marker of malignant strictures, an exploratory study of CA19-9 levels was performed (online supplementary figure 1). By ROC curve analysis, serum CA19-9 could distinguish malignant from benign strictures with an AUC of 0.727. At a threshold of ≥ 44 U/mL chosen using Youden's index, an elevated CA19-9 yielded a sensitivity of 62% and specificity of 77% for detecting malignancy ([table 2](#)). Specimens obtained during ERCP for pathological evaluation included biliary brushings for 22 (39%) patients, biliary biopsies for 22 (39%) patients and both biliary brushings and biopsies for 13 (22%) patients. A pathological diagnosis of at least suspicious for adenocarcinoma was associated with a sensitivity and specificity of 46% and 100%, respectively, for malignancy.

A paired but separate biliary brushing and/or biopsy specimen was also submitted for BiliSeq testing. However, for the 13 patients who had both brushing and biopsy specimens for pathological evaluation, only two patients had paired brushing and 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. Genomic alterations were only detected in malignant strictures and not present in cases with benign cholangiopathy ([figure 1](#)). The sensitivity and specificity of a genomic alteration identified by BiliSeq was 63% and 100%, respectively.

A validation cohort was prospectively collected and consisted of 195 patients with follow-up for 163 (84%) patients ([table 1](#),

online supplementary data and [table 1](#)). A serum CA19-9 was measured at the time of presentation for 144 of 163 (88%) patients. Specimens submitted for pathological evaluation included biliary brushings for 49 (30%) patients, biliary biopsies for 57 (35%) patients and both biliary brushings and biopsies for 57 (35%) patients. A paired but separate brushing and/or biopsy specimen was also submitted for BiliSeq testing. For the 57 patients that had both brushing and biopsy specimens for pathological evaluation, 27 (47%) also had separate brushing and biopsy specimens submitted for BiliSeq testing. Of the remaining 30 patients, corresponding brushings or biopsies were obtained for 24 patients and 6 patients, respectively, for BiliSeq testing. In total, 190 bile duct specimens from 163 patients were submitted for BiliSeq testing.

Based on cut-off of ≥ 44 U/mL, serum CA19-9 within the validation cohort had a sensitivity of 81% and specificity of 65% for at least high-grade biliary dysplasia ([table 2](#)). Similar to the training cohort, the performance of pathological evaluation and BiliSeq testing was associated with sensitivities and specificities of 49% and 98%, and 76% and 100%, respectively. An overall analysis of both training and validation cohorts that included 220 patients with follow-up revealed sensitivities and specificities for serum CA19-9, pathological evaluation and BiliSeq testing of 76% and 69%, 48% and 99%, and 73% and 100%. The combination of pathological evaluation and BiliSeq testing increased the sensitivity of both assays to 83% and maintained a high specificity of 99%. The addition of serum CA19-9 to pathological evaluation and BiliSeq testing further improved the sensitivity to 95%, but the specificity decreased to 68%.

A subanalysis of bile duct specimen subtypes, history of PSC and repeat ERCP with biomarker testing

A total of 132 biliary brushings and 114 biliary biopsies had paired pathological evaluation and BiliSeq testing ([table 3](#)). The sensitivities and specificities for pathological evaluation of biliary brushings and biliary biopsies was 35% and 98%, and 52% and 100%, respectively. In comparison, BiliSeq testing of biliary brushings and biliary biopsies was associated with higher sensitivities of 71% and 75%, respectively, and specificities of 100%. Among 114 biliary biopsies, 28 specimens were obtained by digital cholangioscopy, while the remaining specimens were acquired through cholangiography. No statistically significant

Table 3 Subanalysis of paired biliary brushings, biliary biopsies, distal bile duct strictures, primary sclerosing cholangitis (PSC) and repeat testing

Parameter	Sensitivity, n (95% CI)	Specificity, n (95% CI)	PPV, n (95% CI)	NPV, n (95% CI)	AUC, n (95% CI)
Paired biliary brushings (n=132)					
Pathological evaluation	35 (0.25 to 0.46)	98 (0.88 to 1.00)	97 (0.81 to 1.00)	46 (0.36 to 0.56)	0.66 (0.57 to 0.75)
BiliSeq testing	71 (0.60 to 0.81)	100 (0.91 to 1.00)	100 (0.93 to 1.00)	67 (0.54 to 0.77)	0.86 (0.79 to 0.92)
Pathological evaluation and BiliSeq testing	81 (0.71 to 0.88)	98 (0.88 to 1.00)	99 (0.91 to 1.00)	75 (0.62 to 0.84)	0.89 (0.84 to 0.95)
Paired biliary biopsies (n=114)					
Pathological evaluation	52 (0.41 to 0.62)	100 (0.85 to 1.00)	100 (0.90 to 1.00)	38 (0.27 to 0.51)	0.76 (0.68 to 0.85)
BiliSeq testing	75 (0.65 to 0.84)	100 (0.85 to 1.00)	100 (0.93 to 1.00)	55 (0.40 to 0.70)	0.88 (0.82 to 0.94)
Pathological evaluation and BiliSeq testing	84 (0.74 to 0.91)	100 (0.84 to 1.00)	100 (0.94 to 1.00)	65 (0.48 to 0.79)	0.92 (0.87 to 0.97)
Distal bile duct strictures (n=61)					
Pathological evaluation of biliary specimens	59 (0.42 to 0.73)	95 (0.73 to 1.00)	96 (0.78 to 1.00)	53 (0.36 to 0.69)	0.68 (0.55 to 0.82)
EUS-FNA of biliary stricture	58 (0.42 to 0.73)	100 (0.80 to 1.00)	100 (0.83 to 1.00)	54 (0.37 to 0.70)	0.79 (0.68 to 0.90)
BiliSeq testing	76 (0.59 to 0.87)	100 (0.80 to 1.00)	100 (0.86 to 1.00)	67 (0.47 to 0.82)	0.88 (0.79 to 0.96)
Pathological evaluation (biliary specimens) and BiliSeq testing	90 (0.76 to 0.97)	95 (0.73 to 1.00)	97 (0.85 to 1.00)	83 (0.61 to 0.94)	0.93 (0.85 to 1.00)
Pathological evaluation (biliary and EUS-FNA) and BiliSeq testing	93 (0.79 to 0.98)	95 (0.73 to 1.00)	97 (0.85 to 1.00)	86 (0.64 to 0.96)	0.94 (0.87 to 1.00)
PSC patients (n=37)					
Elevated CA19-9 (n=37)*	67 (0.35 to 0.89)	84 (0.63 to 0.95)	57 (0.30 to 0.81)	83 (0.61 to 0.94)	0.71 (0.53 to 0.90)
Pathological evaluation of biliary brushings/biopsies	8 (0.01 to 0.40)	100 (0.83 to 1.00)	100 (0.06 to 1.00)	69 (0.52 to 0.83)	0.54 (0.34–0.75)
BiliSeq testing	83 (0.51 to 0.98)	100 (0.83 to 1.00)	100 (0.66 to 1.00)	93 (0.74 to 0.99)	0.92 (0.79–1.00)
Elevated CA19-9* and BiliSeq testing	92 (0.60 to 1.00)	84 (0.63 to 0.95)	65 (0.39 to 0.85)	95 (0.73 to 1.00)	0.84 (0.70–0.98)
Pathological evaluation and BiliSeq testing	83 (0.51 to 0.98)	100 (0.83 to 1.00)	100 (0.66 to 1.00)	93 (0.74 to 0.99)	0.92 (0.79 to 1.00)
Elevated CA19-9*, pathological evaluation and BiliSeq testing	92 (0.60 to 1.00)	84 (0.63 to 0.95)	65 (0.39 to 0.85)	95 (0.73 to 1.00)	0.84 (0.70 to 0.98)
Repeat ERCP and biomarker testing (n=220)					
Pathological evaluation of biliary brushings/biopsies	51 (0.43 to 0.60)	99 (0.91 to 1.00)	99 (0.92 to 1.00)	49 (0.40 to 0.57)	0.75 (0.69 to 0.81)
BiliSeq testing	75 (0.68 to 0.82)	100 (0.94 to 1.00)	100 (0.96 to 1.00)	65 (0.56 to 0.74)	0.88 (0.83 to 0.92)
Pathological evaluation and BiliSeq testing	87 (0.80 to 0.92)	99 (0.91 to 1.00)	99 (0.95 to 1.00)	78 (0.67 to 0.85)	0.93 (0.89 to 0.96)

*On the basis of cases where serum CA19-9 was performed concurrently where elevated CA19-9 was defined as >129U/mL.

AUC, area under the curve; CA19-9, carbohydrate antigen 19-9; ERCP, endoscopic retrograde cholangiopancreatography; EUS-FNA, endoscopic ultrasound-fine needle aspiration; NPV, negative predictive value; PPV, positive predictive value; PSC, primary sclerosing cholangitis.

differences in sensitivity and specificity were identified between cholangioscopic-guided and cholangiographic-guided biliary biopsies (online supplementary table 2). Matching biliary brushing and biliary biopsy specimens for both pathological evaluation and BiliSeq testing were submitted for 29 patients. Pathological evaluation of brushings and biopsies, and BiliSeq testing of brushings and biopsies for these cases had sensitivities of 26%, 39%, 70% and 78%, respectively, and specificities of 100%.

EUS-FNA may improve the diagnosis of malignant distal bile duct strictures. In total, 141 patients with follow-up presented with a distal bile duct stricture and a concurrent EUS was performed for 86 patients and FNA for 61 patients. Among the 61 patients with concurrent ERCP and EUS-FNA, the sensitivities and specificities of detecting a malignant distal bile duct stricture based on pathological evaluation of ERCP-obtained biliary specimens and EUS-FNA specimens were 59% and 95%, and 58% and 100%, respectively. In comparison, BiliSeq testing yielded a sensitivity of 76% and specificity of 100%. Furthermore, combining pathological evaluation of ERCP-obtained specimens and BiliSeq testing was associated with a 90% sensitivity and 95% specificity for malignancy. The sensitivity increased to 93% with the addition of EUS-FNA evaluation and the specificity remained at 95%.

Thirty-seven of 220 (17%) patients had a documented history of PSC. This cohort consisted of 12 bile duct strictures with at least high-grade biliary dysplasia and 25 biliary strictures with benign cholangiopathy. As per the American Association for the Study of Liver Diseases (AASLD) guidelines, a serum CA19-9 of >129U/mL

is a marker of suspected cholangiocarcinoma within this high-risk population.⁵⁰ Therefore, based on a cut-off of 129U/mL, serum CA19-9 was associated with a 67% sensitivity and 84% specificity for at least high-grade biliary dysplasia. Pathological evaluation had a sensitivity of only 8% and specificity of 100%. In contrast, the sensitivity and specificity of BiliSeq testing was 83% and 100%, respectively. Combining serum CA19-9 to BiliSeq testing increased the sensitivity to 92% but decreased the specificity to 84%. The addition of pathological evaluation to BiliSeq testing did not improve the sensitivity of BiliSeq alone.

A repeat ERCP with biliary brushings and/or biopsies for pathological evaluation and BiliSeq testing was performed for 20 of 220 (9%) patients (online supplementary table 3). This procedure was repeated once for 16 patients, twice for two patients and thrice for two patients. On the basis of follow-up, 15 patients had a malignant stricture and five patients had a benign stricture. Among the 15 patients with a malignant stricture, 8 patients had genomic alterations detected by BiliSeq on initial and repeat bile duct specimens. However, corresponding pathological evaluation of initial biliary brushings and/or biopsies were negative for malignancy. In fact, a preoperative diagnosis of at least suspicious for malignancy was never rendered for 5 of 8 malignant strictures. For the remaining seven patients, repeat biomarker testing was positive for malignancy by pathological evaluation and detectable genomic alterations ("positive") by BiliSeq testing for two patients; negative by pathology, but positive by BiliSeq for two patients; and both diagnostic modalities were negative for one patient. Inclusion of repeat ERCPs marginally increased the sensitivity and specificities of pathological evaluation, BiliSeq testing and the combination of

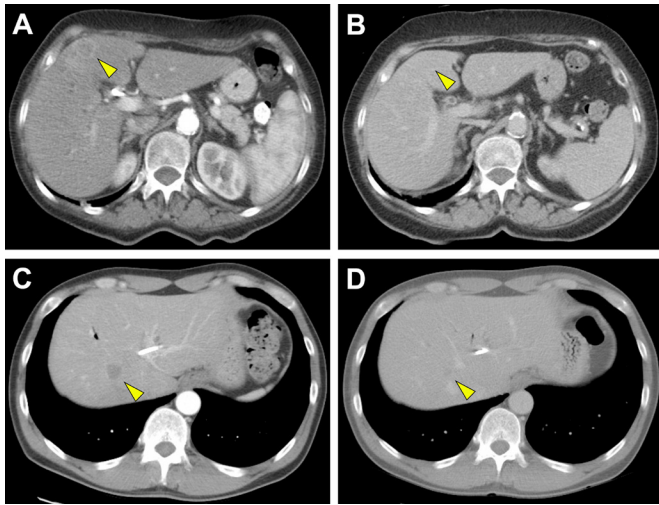


Figure 2 Radiographic response to trastuzumab for two patients with *ERBB2* amplification by BiliSeq testing. Axial CT images of an elderly female (patient 157) with *ERBB2*-amplified metastatic intrahepatic cholangiocarcinoma (A) before and (B) after 1 year with gemcitabine, cisplatin and trastuzumab therapy with decrease in size of the rim-enhancing hypodense hepatic metastasis from 2.1 cm to 0.3 cm. A similar response was seen in a middle age male (patient 166) with 2.6 cm rim-enhancing hypodense hepatic metastasis from a perihilar cholangiocarcinoma (C) that harboured an *ERBB2* amplification. (D) After the combination of first-line therapy and trastuzumab, the patient's metastatic lesion decreased in size to 0.7 cm in 8 months.

both assays to 51% and 99%, 75% and 100%, and 87% and 99%, respectively.

Comparative BiliSeq testing of paired ERCP-obtained bile duct specimens and diagnostic pathology specimens

Diagnostic pathological material was available for 145 of 220 (66%) patients and consisted of 134 malignancies and 11 benign cholangiopathies. BiliSeq testing was performed on the diagnostic pathology specimen for 54 of 134 (40%) malignancies (online supplementary table 4). Genomic alterations were detected in all 54 cases that also included 12 patients with ERCP-obtained bile duct specimens that were negative for genomic alterations by BiliSeq. Among the remaining 42 tumours, the genomic profile was different for four cases that included additional genomic alterations than those detected by BiliSeq using ERCP-obtained bile duct specimens.

Therapeutically relevant genomic alterations and treatment data

On the basis of previously reported actionable targets and predictive markers of chemotherapeutic response, BiliSeq testing of ERCP-obtained bile duct specimens identified therapeutically relevant genomic alterations for 20 of 150 (13%) cancers (online supplementary figure 2). Alterations in genes involved in the receptor tyrosine kinase (RTK)/Ras/Mitogen-activated protein (MAP) kinase signalling pathway, which included *ERBB2* (n=6), *MET* (n=4), *ALK* (n=2), *FGFR2* (n=3) and *FGFR3* (n=2), were detected in the majority of cases (n=13, 65%). Other potentially actionable alterations were found in *ATM* (n=2), *IDH1* (n=1) and *IDH2* (n=2).

Among the *ERBB2*-altered malignancies, two patients, whose adenocarcinoma harboured a copy number gain in *ERBB2*, received targeted chemotherapy to HER2/neu instead of standard chemotherapy alone. The first patient (online

supplementary table 1, patient 157) is an elderly woman presenting with jaundice, worsening abdominal pain for a week, elevated liver function tests, a serum CA19-9 of 55 U/mL and a CT scan that showed gallstones and intrahepatic ductal dilatation. An ERCP demonstrated an intrahepatic bile duct stricture that was brushed for both pathological evaluation and BiliSeq testing. Pathological evaluation of the biliary brushings revealed highly atypical epithelial cells, suspicious for adenocarcinoma. However, BiliSeq testing detected an *ERBB2* copy number gain and a *TP53* mutation (p.C242fs*5 and c.723delC). Within 2 weeks, a subsequent triphasic CT scan revealed a 2.1 cm metastatic liver lesion in segment 4B/5 (figure 2A), and her serum CA19-9 was 422.4 U/mL. A diagnosis of malignancy was confirmed after repeat ERCP with both bile duct brushings and biopsy as an invasive moderately-differentiated adenocarcinoma. Clinically considered to be an intrahepatic cholangiocarcinoma, the patient was started on gemcitabine, cisplatin and trastuzumab. Within 3 months of treatment, her serum CA19-9 decreased to 14 U/mL and the metastatic liver lesion decreased to 0.6 cm. Furthermore, a follow-up ERCP noted an improvement in the patient's bile duct stricture (online supplementary figure 3). One year after diagnosis, the patient is currently alive and well. Her serum CA19-9 continues to be within normal limits, and the metastatic lesion measures 0.3 cm in greatest dimension (figure 2B).

The second patient (online supplementary table 1, patient 166) is a middle-aged man who presented with jaundice, upper abdominal pain, elevated liver function tests, an elevated serum CA19-9 of 2383 U/mL, thickening of his perihilar bile duct on CT and a hilar stricture on MRCP. A Bismuth IIIa stricture was identified on ERCP and bile duct brushings and biopsy were positive for adenocarcinoma. The patient's corresponding BiliSeq testing revealed a copy number gain in *ERBB2* and copy number losses in *TP53* and *CDKN2A*. The patient was referred for a liver transplantation, but on further workup was found to have a 2.6 cm metastatic lesion within segment 7/8 (figure 2C). At a serum CA19-9 of 7959 U/mL, the patient started a regimen of gemcitabine, cisplatin and trastuzumab for a clinical perihilar cholangiocarcinoma. After 3 months, the patient's serum CA19-9 normalised to 9.4 U/mL. Eight months after diagnosis, the patient is currently alive and well. His serum CA19-9 continues to be within normal limits and the metastatic lesion decreased to 0.7 cm in greatest dimension (figure 2D).

DISCUSSION

Despite improvements in cross-sectional imaging, serum studies, endoscopic technologies and pathological evaluation, the distinction between benign and malignant bile duct strictures remains a diagnostic conundrum. In the current study, we developed and independently clinically validated a targeted NGS assay for the evaluation of bile duct strictures. We found our targeted NGS assay, BiliSeq, had a higher sensitivity (73% vs 48%) and a higher specificity (100% vs 99%) than pathological evaluation alone for detecting at least high-grade biliary dysplasia involving the bile duct. Moreover, the combination of both tests (BiliSeq and pathological evaluation) had a higher sensitivity than either test alone (83%), while maintaining a high specificity (99%).

To date, a few studies have investigated the application of targeted NGS to bile duct strictures. Recently, Bankov *et al* published the results of an NGS assay of similar design to BiliSeq using a retrospective cohort of 16 patients with indeterminate bile duct biopsies and cholangiocarcinoma on follow-up.⁵¹ The authors did not evaluate biliary brushing specimens and DNA

obtained for NGS was extracted from FFPE tissue blocks. The authors acknowledged that due to the low input and low quality of DNA from FFPE tissue, a DNA enrichment technique was required for subsequent analysis. While a benign control group was not used for comparison, the sensitivity of their assay was 81%, which is comparable with BiliSeq. A more comprehensive assessment of NGS testing on bile duct specimens was reported by Dudley *et al.*⁶ The authors evaluated a cohort of 73 bile duct and 8 main pancreatic duct brushing specimens. Contrary to Bankov *et al.*, their study cohort did not include bile duct biopsies. Furthermore, rather than obtaining a dedicated brushing for DNA isolation, DNA was extracted from CytoLyt-preserved specimens. As alcohol fixation may result in DNA degradation, it is not unexpected that 8 of 73 (11%) brushing specimens failed NGS testing. Within our study, none of the biliary brushing and/or biopsy specimens failed BiliSeq testing. Regardless, among the remaining 65 biliary brushing specimens, Dudley *et al.* reported a sensitivity of 68% and a specificity of 97%, which was superior to pathological evaluation. The authors did however identify one false positive result. For this case, the duration of follow-up after NGS testing was unclear and, per the authors, the possibility of dysplasia within the bile duct system could not be excluded.

Although BiliSeq did not identify any false positive cases, there were 37 (25%) false-negative cases after accounting for repeat testing. In order to determine the cause of false-negative NGS results, we performed BiliSeq testing on available diagnostic pathological material from 12 of these 37 malignancies. A genomic alteration in at least one of the 28 genes was detected within our molecular panel. Considering the collection of bile duct specimens is technically difficult and sampling failures are known to occur, inadequate sampling of the stricture may be a possible explanation. Alternatively, low specimen tumour cellularity may also be a factor. The lowest limit of detection for our NGS assay is approximately 3% of mutant alleles or, for a heterozygous mutation, such as *KRAS*, 6% tumour cellularity. Therefore, if the fraction of tumour cells was lower than 6%, then BiliSeq would be unable to detect a pathogenic genomic alteration. While this finding could be viewed as a technical constraint, controlling the sensitivity of BiliSeq may be beneficial, especially for specific genes. Previous studies using surgical resection material have demonstrated *KRAS* mutations in cases of low-grade biliary dysplasia and benign cholangiopathy.⁵² Therefore, increasing the sensitivity of BiliSeq for *KRAS* mutations may result in a decrease in specificity for at least high-grade biliary dysplasia. It is also important to emphasise that a subset of malignant neoplasms involving the bile duct may harbour genomic alterations that are not assayed by our test. For example, mutations in chromatin remodelling genes (eg, *ARID1A*) and *FGFR2* fusion genes are present in up to 20% of intrahepatic cholangiocarcinomas.^{26 30 32–34} Thus, the sensitivity of BiliSeq could potentially be improved with target amplification/enrichment techniques and an expanded gene panel.

Another intriguing, but preliminary observation, from our study is the striking difference in sensitivity between BiliSeq and pathological evaluation for at least high-grade biliary dysplasia among patients with PSC. PSC is an autoimmune condition and associated with a lifetime incidence of cholangiocarcinoma that ranges from 6% to 36%.^{53–56} The risk of cholangiocarcinoma is highest within the first 2 years of a PSC diagnosis and suggests a possible window of opportunity for early detection.^{55 57} However, a small cholangiocarcinoma in a background of an inflammatory stricture is particularly difficult to sample and challenging to diagnose by pathological evaluation. Although the number of patients with PSC within our cohort was relatively

small, the sensitivity of pathological evaluation was only 8%. In comparison, BiliSeq achieved an 83% sensitivity for at least high-grade biliary dysplasia and provides compelling evidence to further explore its potential to improve clinical outcomes for this high-risk patient population.

Besides NGS testing, several adjunctive techniques have been developed clinically to evaluate bile duct strictures, such as digital image analysis, single gene sequencing and multifocal FISH for chromosomal polysomy. Among these ancillary studies, multicolour FISH has become an adopted tool by many academic institutions for over a decade.¹⁵ Analogous to BiliSeq, the premise of multicolour FISH is the finding that dysplastic and malignant neoplasms involving the bile duct are often characterised by a high frequency of numerical chromosomal abnormalities. However, this test typically achieves sensitivities of only 35%–60% with minimal technological improvement in the past few years.^{13 15 18} Multicolour FISH is expensive, labour intensive, prone to subjective interpretation errors and requires significant technical expertise. In fact, the interpretation of multicolour FISH requires an experienced pathologist, since a paired morphological assessment of the specimen is needed to ensure that the epithelial cells of interest are being evaluated for chromosomal abnormalities. While we did not perform a formal comparison between BiliSeq and multicolour FISH, Dudley *et al.*⁶ found their NGS assay had a higher sensitivity than multicolour FISH. This is not surprising as NGS testing detects copy number alterations and single nucleotide variants in multiple genes. Furthermore, with decreasing costs and the ability to batch multiple specimens in a single run, NGS testing has become widely available to both academic and non-academic institutions. Thus, NGS testing may be a more preferable option than multicolour FISH in the evaluation of bile duct specimens. Demonstrating the clinical feasibility of BiliSeq to ERCP-obtained biliary specimens also lays the foundation for further applications of NGS testing to non-invasive specimen types, such as bile and serum (eg, circulating tumour DNA and exosomes).^{58–61}

In addition to early detection, a major impetus for widespread adoption of NGS technologies has been solid tumour genotyping. Malignant neoplasms arising or secondarily involving the bile duct system are typically recalcitrant and current standard-of-care chemotherapeutic regimens are associated with limited efficacy. On relapse or progression after first-line or second-line chemotherapy, comprehensive genomic profiling for a subset of tumours has identified alterations that are potentially targetable or predictive markers of response to specific therapy.^{26 28} For example, *ERBB2* amplification has been reported in intrahepatic cholangiocarcinoma, perihilar/distal cholangiocarcinoma and gallbladder adenocarcinoma with a prevalence of 5%, 17% and 19%, respectively.^{28 62} Among eight patients with *ERBB2*-amplified or *HER2/neu*-overexpressed gallbladder adenocarcinoma, Javle *et al.*⁴⁶ reported three had stable disease, four had partial responses and one had a complete response to targeted therapy (trastuzumab or trastuzumab with pertuzumab). Of note, five of the eight patients received first-line chemotherapy or chemoradiation prior to treatment with targeted therapy. Within our study, BiliSeq testing identified two patients with *ERBB2* amplification: one intrahepatic cholangiocarcinoma and one perihilar cholangiocarcinoma. Both patients received trastuzumab in conjunction with standard first-line chemotherapy (gemcitabine and cisplatin), both exhibited measurable radiographic response and normalisation of serum CA19-9, and both are currently alive and well. While we recognise the benefits of standard chemotherapy, it seems reasonable to assume the

inclusion of trastuzumab within the chemotherapeutic regimen to be responsible for the significant and prolonged treatment effects observed in both patients. It should also be stressed that the identification of a targetable alteration at the time of diagnosis facilitated a personalised approach to treatment. Moreover, *ERBB2* was not the only genomic alteration detected by BiliSeq to potentially confer susceptibility to anticancer therapy.

We acknowledge that there are several limitations to this study. Although a large number of bile duct specimens were analysed, follow-up diagnostic pathology was available for only 66% of patients. However, the majority of patients with malignant neoplasms involving the bile duct system often have unresectable disease at presentation or are poor surgical candidates. In addition, in order to ensure sufficient DNA for NGS testing, a separate bile duct specimen than the specimen used for pathological evaluation was used for BiliSeq testing. This protocol may explain the incongruent results between pathological evaluation and BiliSeq testing for a subset of cases. Furthermore, this study does not address the optimal gene panel or approach of integrating NGS-based molecular testing to the evaluation of a bile duct stricture. To minimise the need for repeat sampling, we routinely perform BiliSeq on all patients with potentially malignant strictures at the time of initial ERCP. A positive BiliSeq result hastens the referral process for appropriate surgical and/or oncological subspecialists. A negative BiliSeq result provides additional reassurance to negative pathological evaluation. Because none of these diagnostic modalities achieves a sensitivity of >95%, few patients can be fully dismissed from surveillance based on a negative BiliSeq. A similar approach has been adopted by investigators using multicolour FISH. It is also important to recognise that recent advancements in cholangioscopy and increasing use of EUS-FNA may potentially improve the detection of malignant strictures.^{63,64} While a minority of patients within this study were evaluated by cholangioscopy, we found similar diagnostic performance characteristics between cholangioscopy and cholangiography.⁶⁵ The application of EUS-FNA to bile duct strictures, especially distal strictures, is reported to enhance tissue acquisition over traditional ERCP methods.⁶⁶ However, patients in our study cohort evaluated using both techniques did not show any statistically significant differences in sensitivity. In fact, BiliSeq yielded a higher sensitivity than ERCP or EUS-FNA alone and, in combination with either method, increased the sensitivity of detecting a distal malignant stricture to at least 90%. Lastly, our definition of targetable genomic alterations is based on an aggregation of a myriad of studies. The true clinical utility of any given drug to a particular genomic alteration is unknown until a prospective clinical trial is performed. However, we were able to demonstrate at least anecdotal evidence of clinical benefit by pairing genomic alterations, such as *ERBB2* amplification, with specific anticancer therapy (eg, trastuzumab).

In summary, we report the largest prospective study to examine the role of NGS-based molecular (BiliSeq) testing to the evaluation of bile duct strictures. Our results support the clinical utility of combining BiliSeq to standard pathological evaluation of bile duct specimens. Additionally, BiliSeq has the potential to identify targetable genomic alterations and, therefore, stratify patients for specific chemotherapeutic regimens. Future studies are however required to explore the integration of BiliSeq and similar NGS-based assays into current surveillance and management guidelines for patients with bile duct strictures.

Author affiliations

¹Department of Pathology, University of Pittsburgh Medical Center Health System, Pittsburgh, Pennsylvania, USA

²Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

³Department of Gastroenterology, Hepatology, and Nutrition, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA

⁴Department of Pathology, Division of Transplant Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

⁵Department of Surgery, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

⁶Department of Clinical Sciences, Surgery, University of Texas Southwestern, Dallas, Texas, USA

⁷Department of Surgery, West Virginia University Health Sciences Center, Morgantown, West Virginia, USA

⁸Department of Surgery, NorthShore University Health System, Evanston, Illinois, USA

⁹Department of Surgery, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA

¹⁰Department of Transplant, Thomas E Starzl Transplant Institute University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA

¹¹Division of Hematology and Oncology, UPMC Cancer Centers, Pittsburgh, Pennsylvania, USA

¹²Department of Radiology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Acknowledgements The authors would like to thank Mrs Kate Smith for outstanding administrative assistance. This study was supported in part by the Institute for Precision Medicine at the University of Pittsburgh, the UPMC Hillman Cancer Center, Shear Family Foundation, the Pittsburgh Liver Research Center at the University of Pittsburgh and the University of Pittsburgh Medical Center, and the Sky Foundation (to ADS).

Contributors Study concept and design: ADS, MNN and AS. Acquisition of data: all authors. Analysis and interpretation of data: ADS, MNN and AS. Drafting of the manuscript: ADS and AS.

Funding This study was funded by the University of Pittsburgh Institute of Precision Medicine.

Competing interests ADS has received an honorarium from Foundation Medicine, Inc.

Patient consent for publication Obtained.

Ethics approval Study approval was obtained from the University of Pittsburgh Institutional Review Board (IRB# PRO17030748).

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Aatur D Singhi <http://orcid.org/0000-0003-3930-7096>

REFERENCES

- Singh A, Gelrud A, Agarwal B. Biliary strictures: diagnostic considerations and approach. *Gastroenterol Rep* 2015;3:22–31.
- Nguyen Canh H, Harada K. Adult bile duct strictures: differentiating benign biliary stenosis from cholangiocarcinoma. *Med Mol Morphol* 2016;49:189–202.
- Anon. NIH state-of-the-science statement on endoscopic retrograde cholangiopancreatography (ERCP) for diagnosis and therapy. *NIH Consens State Sci Statements* 2002;19:1–26.
- Korc P, Sherman S. ERCP tissue sampling. *Gastrointest Endosc* 2016;84:557–71.
- Chen YK, Parsi MA, Binmoeller KF, et al. Single-operator cholangioscopy in patients requiring evaluation of bile duct disease or therapy of biliary stones (with videos). *Gastrointest Endosc* 2011;74:805–14.
- Dudley JC, Zheng Z, McDonald T, et al. Next-Generation Sequencing and Fluorescence in Situ Hybridization Have Comparable Performance Characteristics in the Analysis of Pancreaticobiliary Brushings for Malignancy. *J Mol Diagn* 2016;18:124–30.
- Foutch PG, Kerr DM, Harlan JR, et al. A prospective, controlled analysis of endoscopic cytotechniques for diagnosis of malignant biliary strictures. *Am J Gastroenterol* 1991;86:577–80.
- Farrell RJ, Jain AK, Brandwein SL, et al. The combination of stricture dilation, endoscopic needle aspiration, and biliary brushings significantly improves diagnostic yield from malignant bile duct strictures. *Gastrointest Endosc* 2001;54:587–94.
- Kurzawinski TR, Deery A, Dooley JS, et al. A prospective study of biliary cytology in 100 patients with bile duct strictures. *Hepatology* 1993;18:1399–403.
- de Bellis M, Fogel EL, Sherman S, et al. Influence of stricture dilation and repeat brushing on the cancer detection rate of brush cytology in the evaluation of malignant biliary obstruction. *Gastrointest Endosc* 2003;58:176–82.
- Stewart CJ, Mills PR, Carter R, et al. Brush cytology in the assessment of pancreaticobiliary strictures: a review of 406 cases. *J Clin Pathol* 2001;54:449–55.

- 12 Levy MJ, Baron TH, Clayton AC, *et al.* Prospective evaluation of advanced molecular markers and imaging techniques in patients with indeterminate bile duct strictures. *Am J Gastroenterol* 2008;103:1263–73.
- 13 Kipp BR, Fritcher EG, Clayton AC, *et al.* Comparison of KRAS mutation analysis and FISH for detecting pancreatobiliary tract cancer in cytology specimens collected during endoscopic retrograde cholangiopancreatography. *J Mol Diagn* 2010;12:780–6.
- 14 Moreno Luna LE, Kipp B, Halling KC, *et al.* Advanced cytologic techniques for the detection of malignant pancreatobiliary strictures. *Gastroenterology* 2006;131:1064–72.
- 15 Kipp BR, Stadheim LM, Halling SA, *et al.* A comparison of routine cytology and fluorescence in situ hybridization for the detection of malignant bile duct strictures. *Am J Gastroenterol* 2004;99:1675–81.
- 16 Gonda TA, Viterbo D, Gausman V, *et al.* Mutation Profile and Fluorescence In Situ Hybridization Analyses Increase Detection of Malignancies in Biliary Strictures. *Clin Gastroenterol Hepatol* 2017;15:913–9.
- 17 Brooks C, Gausman V, Kokoy-Mondragon C, *et al.* Role of Fluorescent In Situ Hybridization, Cholangioscopic Biopsies, and EUS-FNA in the Evaluation of Biliary Strictures. *Dig Dis Sci* 2018;63:636–44.
- 18 Barr Fritcher EG, Voss JS, Brankley SM, *et al.* An Optimized Set of Fluorescence In Situ Hybridization Probes for Detection of Pancreatobiliary Tract Cancer in Cytology Brush Samples. *Gastroenterology* 2015;149:1813–24.
- 19 Clayton RA, Clarke DL, Currie EJ, *et al.* Incidence of benign pathology in patients undergoing hepatic resection for suspected malignancy. *Surgeon* 2003;1:32–8.
- 20 Wakai T, Shirai Y, Sakata J, *et al.* Clinicopathological features of benign biliary strictures masquerading as biliary malignancy. *Am Surg* 2012;78:1388–91.
- 21 Thelen A, Neuhaus P. Liver transplantation for hilar cholangiocarcinoma. *J Hepatobiliary Pancreat Surg* 2007;14:469–75.
- 22 Neuhaus P, Jonas S. Surgery for hilar cholangiocarcinoma—the German experience. *J Hepatobiliary Pancreat Surg* 2000;7:142–7.
- 23 Neuhaus P, Jonas S, Bechstein WO, *et al.* Extended resections for hilar cholangiocarcinoma. *Ann Surg* 1999;230:808–18. discussion 819.
- 24 Kosuge T, Yamamoto J, Shimada K, *et al.* Improved surgical results for hilar cholangiocarcinoma with procedures including major hepatic resection. *Ann Surg* 1999;230:663–71.
- 25 Jiao Y, Pawlik TM, Anders RA, *et al.* Exome sequencing identifies frequent inactivating mutations in BAP1, ARID1A and PBRM1 in intrahepatic cholangiocarcinomas. *Nat Genet* 2013;45:1470–3.
- 26 Lowery MA, Ptashkin R, Jordan E, *et al.* Comprehensive Molecular Profiling of Intrahepatic and Extrahepatic Cholangiocarcinomas: Potential Targets for Intervention. *Clin Cancer Res* 2018;24:4154–61.
- 27 Lowery MA, Jordan EJ, Basturk O, *et al.* Real-Time Genomic Profiling of Pancreatic Ductal Adenocarcinoma: Potential Actionability and Correlation with Clinical Phenotype. *Clin Cancer Res* 2017;23:6094–100.
- 28 Javle M, Bekaii-Saab T, Jain A, *et al.* Biliary cancer: Utility of next-generation sequencing for clinical management. *Cancer* 2016;122:3838–47.
- 29 Bailey P, Chang DK, Nones K, *et al.* Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 2016;531:47–52.
- 30 Farshidfar F, Zheng S, Gingras MC, *et al.* Integrative Genomic Analysis of Cholangiocarcinoma Identifies Distinct IDH-Mutant Molecular Profiles. *Cell Rep* 2017;18:2780–94.
- 31 Nakamura H, Arai Y, Totoki Y, *et al.* Genomic spectra of biliary tract cancer. *Nat Genet* 2015;47:1003–10.
- 32 Nepal C, O'Rourke CJ, Oliveira D, *et al.* Genomic perturbations reveal distinct regulatory networks in intrahepatic cholangiocarcinoma. *Hepatology* 2018;68:949–63.
- 33 Jusakul A, Cutcutache I, Yong CH, *et al.* Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. *Cancer Discov* 2017;7:1116–35.
- 34 Chaisaingmongkol J, Budhu A, Dang H, *et al.* Common Molecular Subtypes Among Asian Hepatocellular Carcinoma and Cholangiocarcinoma. *Cancer Cell* 2017;32:57–70.
- 35 Singhi AD, McGrath K, Brand RE, *et al.* Preoperative next-generation sequencing of pancreatic cyst fluid is highly accurate in cyst classification and detection of advanced neoplasia. *Gut* 2018;67:2131–41.
- 36 Nikiforova MN, Mercurio S, Wald AI, *et al.* Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. *Cancer* 2018;124:1682–90.
- 37 Singhi AD, George B, Greenbowe JR, *et al.* Real-time Targeted Genome Profile Analysis of Pancreatic Ductal Adenocarcinomas Identifies Genetic Alterations that Might be Targeted with Existing Drugs or Used as Biomarkers. *Gastroenterology* 2019.
- 38 Bosman FT, Carneiro F, Hruban RH, *et al.* *World Health Organization (WHO) Classification of Tumours of the Digestive System*. Lyon, France: IARC Press, 2010.
- 39 Nikiforova MN, Wald AI, Roy S, *et al.* Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J Clin Endocrinol Metab* 2013;98:E1852–60.
- 40 Li MM, Datto M, Duncavage EJ, *et al.* Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn* 2017;19:4–23.
- 41 Grasso C, Butler T, Rhodes K, *et al.* Assessing copy number alterations in targeted, amplicon-based next-generation sequencing data. *J Mol Diagn* 2015;17:53–63.
- 42 Nikiforova MN, Wald AI, Melan MA, *et al.* Targeted next-generation sequencing panel (GlioSeq) provides comprehensive genetic profiling of central nervous system tumors. *Neuro Oncol* 2016;18:379–87.
- 43 Javle M, Lowery M, Shroff RT, *et al.* Phase II Study of BGI398 in Patients With FGFR-Altered Advanced Cholangiocarcinoma. *J Clin Oncol* 2018;36:276–82.
- 44 Chen SH, Zhang Y, Van Horn RD, *et al.* Oncogenic BRAF Deletions That Function as Homodimers and Are Sensitive to Inhibition by RAF Dimer Inhibitor LY3009120. *Cancer Discov* 2016;6:300–15.
- 45 Singhi AD, Ali SM, Lacy J, *et al.* Identification of Targetable ALK Rearrangements in Pancreatic Ductal Adenocarcinoma. *J Natl Compr Canc Netw* 2017;15:555–62.
- 46 Javle M, Churi C, Kang HC, *et al.* HER2/neu-directed therapy for biliary tract cancer. *J Hematol Oncol* 2015;8:58.
- 47 Kaufman B, Shapira-Frommer R, Schmutzler RK, *et al.* Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. *J Clin Oncol* 2015;33:244–50.
- 48 Rohle D, Popovici-Muller J, Palaskas N, *et al.* An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. *Science* 2013;340:626–30.
- 49 Saha SK, Gordan JD, Kleinstiver BP, *et al.* Icositate Dehydrogenase Mutations Confer Dasatinib Hypersensitivity and SRC Dependence in Intrahepatic Cholangiocarcinoma. *Cancer Discov* 2016;6:727–39.
- 50 Chapman R, Fevery J, Kalloo A, *et al.* Diagnosis and management of primary sclerosing cholangitis. *Hepatology* 2010;51:660–78.
- 51 Bankov K, Döring C, Schneider M, *et al.* Sequencing of intraductal biopsies is feasible and potentially impacts clinical management of patients with indeterminate biliary stricture and cholangiocarcinoma. *Clin Transl Gastroenterol* 2018;9:e151.
- 52 Hsu M, Sasaki M, Igarashi S, *et al.* KRAS and GNAS mutations and p53 overexpression in biliary intraepithelial neoplasia and intrahepatic cholangiocarcinomas. *Cancer* 2013;119:1669–74.
- 53 Patel T. Increasing incidence and mortality of primary intrahepatic cholangiocarcinoma in the United States. *Hepatology* 2001;33:1353–7.
- 54 Boonstra K, Weersma RK, van Erpecum KJ, *et al.* Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 2013;58:2045–55.
- 55 Claessen MM, Vleggaar FP, Tytgat KM, *et al.* High lifetime risk of cancer in primary sclerosing cholangitis. *J Hepatol* 2009;50:158–64.
- 56 LaRusso NF, Sneider BL, Black D, *et al.* Primary sclerosing cholangitis: summary of a workshop. *Hepatology* 2006;44:746–64.
- 57 Burak K, Angulo P, Pasha TM, *et al.* Incidence and risk factors for cholangiocarcinoma in primary sclerosing cholangitis. *Am J Gastroenterol* 2004;99:523–6.
- 58 Andersen RF, Jakobsen A. Screening for circulating RAS/RAF mutations by multiplex digital PCR. *Clin Chim Acta* 2016;458:138–43.
- 59 Li L, Masica D, Ishida M, *et al.* Human bile contains microRNA-laden extracellular vesicles that can be used for cholangiocarcinoma diagnosis. *Hepatology* 2014;60:896–907.
- 60 Arbelaez A, Azkargorta M, Krawczyk M, *et al.* Serum extracellular vesicles contain protein biomarkers for primary sclerosing cholangitis and cholangiocarcinoma. *Hepatology* 2017;66:1125–43.
- 61 Bernard V, Kim DU, San Lucas FA, *et al.* Circulating Nucleic Acids Are Associated With Outcomes of Patients With Pancreatic Cancer. *Gastroenterology* 2019;156:108–18.
- 62 Galdy S, Lamarca A, McNamara MG, *et al.* HER2/HER3 pathway in biliary tract malignancies; systematic review and meta-analysis: a potential therapeutic target? *Cancer Metastasis Rev* 2017;36:141–57.
- 63 Navaneethan U, Hasan MK, Louridasamy V, *et al.* Single-operator cholangioscopy and targeted biopsies in the diagnosis of indeterminate biliary strictures: a systematic review. *Gastrointest Endosc* 2015;82:608–14.
- 64 Lee JH, Salem R, Aslanian H, *et al.* Endoscopic ultrasound and fine-needle aspiration of unexplained bile duct strictures. *Am J Gastroenterol* 2004;99:1069–73.
- 65 Hartman DJ, Krasinskas AM, Uttam S, *et al.* Assessment of nuclear nanomorphology marker to improve the detection of malignancy from bile duct biopsy specimens. *Am J Clin Pathol* 2014;141:884–91.
- 66 Weiler F, Bhat YM, Binmoeller KF, *et al.* EUS-FNA is superior to ERCP-based tissue sampling in suspected malignant biliary obstruction: results of a prospective, single-blind, comparative study. *Gastrointest Endosc* 2014;80:97–104.