ORIGINAL ARTICLE

Risk factors to differentiate between benign proximal biliary strictures and perihilar cholangiocarcinoma

Fabio Tirotta[1], Francesco Giovinazzo[1], James Hodson[2], Louise Yates[1], Dominic Tan[1], Marco Salsano[3], Diya Baker[1], Ramanivas Sundareyan[3], Ravi Marudanayagam[1], Bobby Dasari[1], Keith Roberts[1], Paolo Muiesan[1], John Isaac[1], Darius Mirza[1], Homoyoon Mehrzad[3] & Robert Sutcliffe[1]

[1]Department of HPB Surgery, [2]Department of Statistics, Institute of Translational Medicine, and [3]Department of Radiology, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

Abstract

Background: The aim of this study was to evaluate potential risk factors associated with benign lesions and perihilar cholangiocarcinoma (PHC) in patients presenting with proximal biliary strictures (PBS).

Methods: Patients with PBS who were referred to a specialist HPB centre between 2008 and 2016 were identified. Patients with primary sclerosing cholangitis, metastatic PHC or hilar obstruction by a peripheral tumour were excluded. The final diagnosis was determined either by (1) resection histology or (2) combination of biopsy and clinical course. Multivariable analysis of clinical, laboratory and radiological data was undertaken to identify independent predictors of benign and malignant lesions.

Results: 155 consecutive patients were identified, including 25 patients (16%) with benign PBS. Abdominal pain (odds ratio [OR] 3.36; p = 0.027), serum CA19.9 < 100 U/ml (OR 10.35; p = 0.001), and absence of mass on imaging (OR 4.66; p = 0.004) were all associated with the presence of benign lesions on multivariable analysis.

Conclusions: This study has identified several independent variables that may differentiate between benign and malignant proximal biliary strictures. A larger multi-institutional study would be warranted to validate these findings, and to develop a risk score to stratify patients with suspected PHC.

Received 8 December 2019; accepted 30 March 2020

Correspondence

Robert Sutcliffe, Department of HPB Surgery, University Hospitals Birmingham NHS Foundation Trust, Mindelsohn Way, Birmingham, B15 2GW, England, UK. E-mail: Robert.Sutcliffe@uhb.nhs.uk

Introduction

Perihilar cholangiocarcinoma (PHC) is a rare malignancy that affects the hepatic duct confluence in the hepatic hilum.[1] Patients typically present with jaundice due to a proximal biliary stricture (PBS) and in the absence of metastatic disease, surgical resection is the only potentially curative treatment. Although the majority of PBS are secondary to PHC, a significant proportion of patients will have benign pathology, such as inflammatory strictures, sclerosing cholangitis or IgG4-related cholangiopathy.[2–4] The diagnostic work-up of patients who present with PBS includes cross-sectional imaging (contrast-enhanced CT and/or MRI) and cholangiography (MRCP, ERCP or PTC) to confirm the diagnosis, assess resectability and exclude metastatic disease. However, the diagnostic specificity of cross-sectional radiology is poor, and differentiating benign and malignant PBS on imaging criteria alone is challenging.[5] Indeed, up to 15% of patients who undergo major hepatectomy for suspected PHC will be found to have benign disease on histopathological analysis.[2,4] Although surgical resection is widely recommended for patients with suspected PHC in the absence of preoperative histological confirmation[5,6] preoperative biopsy may yield diagnostic information in up to 70% of patients.[7] False negative biopsies are common, and repeat attempts at obtaining a preoperative tissue diagnosis may delay potentially curative surgery in patients with PHC. Conversely, patients with benign PBS are over-treated by major hepatectomy and unnecessarily exposed to significant perioperative risks. There is clearly a need to improve the diagnostic pathway of patients with suspected PHC. The aim of this study was to evaluate risk factors for benign disease in patients presenting with PBS.

Methods

Preoperative work-up

All patients with PBS were discussed at a regional Hepatobiliary Multidisciplinary Team (MDT) meeting. Patients suspected to have potentially resectable PHC underwent percutaneous transhepatic biliary drainage with or without portal vein embolization (based on future remnant CT volumetry) prior to surgery. Preoperative biopsy was undertaken selectively, depending on the degree of suspicion for malignancy. Patients with either suspicious imaging or a positive biopsy for malignancy underwent surgery. Patients with neither suspicious imaging nor positive biopsy underwent repeat biopsy and/or interval cross-sectional imaging, according to the MDT opinion.

Data collection

All consecutive patients with PBS who had been referred to the specialist MDT meeting at University Hospital Birmingham between 2008 and 2016 were identified from a prospectively maintained database. Clinical, laboratory and radiological data were collected from the time of clinical presentation for all patients by review of electronic records. Exclusion criteria were (1) extrinsic hilar obstruction due to a peripheral tumour; (2) history of previous malignancy; (3) metastatic PHC; (4) history of (or imaging suggestive of) primary sclerosing cholangitis; (5) post-cholecystectomy stricture; (6) intrahepatic or distal biliary stricture; (7) common bile duct stone and (8) intraductal papillary neoplasm. Cross-sectional (contrast-enhanced CT and/or MRI) and cholangiography (percutaneous or MRCP) images were prospectively reviewed by a specialist hepatobiliary radiologist who was blinded to the final diagnosis. The following radiological features were evaluated: (1) number and length of strictures, (2) presence of a mass, (3) duct/lesion enhancement and (4) ductal separation. The final diagnosis of a malignant lesion was confirmed by biopsy and/or resection histology. The final diagnosis of a benign lesion was determined either by resection histology (operated patients) or after a minimum of 12 months of clinical and imaging follow-up (non-operated patients).

Statistical methods

Clinical, biochemical and radiological features were compared between those patients with final diagnoses of benign vs. malignant strictures. Continuous variables were reported as medians and interquartile ranges (IQRs), with p-values from Mann–Whitney U tests. Ordinal variables were also analysed using Mann–Whitney U test, with Fisher's exact test used for nominal factors. A multivariable analysis was then performed, using a binary logistic regression model with a backwards stepwise approach, in order to identify significant independent predictors of benign strictures. Continuous variables with large quantities of missing data were dichotomised prior to analysis, and an additional "unknown" category was included to minimise exclusions due to missing data. All analyses were performed using IBM SPSS 22 (IBM Corp. Armonk, NY), with $p < 0.05$ deemed to be indicative of statistical significance throughout.

Results

During the study period, 257 patients with suspected PBS were referred to the specialist MDT. 102 patients were excluded due to extrinsic hilar obstruction (30), history of previous malignancy (2), metastatic PHC (27), primary sclerosing cholangitis (21), post-cholecystectomy stricture (1), intrahepatic or distal biliary stricture (11), common bile duct stone (1), intraductal papillary neoplasm (1), imaging not available (3) or no final diagnosis (5). After these exclusions, the study cohort consisted of 155 patients, who had a median age at referral of 68 years (IQR 61–73), of whom 55% were male. The majority of patients (112; 72%) presented without abdominal pain and 123 patients (79%) were clinically jaundiced. Median serum bilirubin and CA19.9 levels at referral were 150 μmol/l (IQR: 41–272) and 253 U/ml (IQR: 65–970), respectively.

Overall, 55 patients (35%) were considered to have either suspected (based on clinical/radiological criteria; N = 27) or biopsy-proven (N = 28) resectable PHC and underwent surgery. 20 patients (74%) with suspected PHC did not undergo a preoperative biopsy, and 7 patients (26%) operated for suspected PHC had a final diagnosis of benign disease. The proportion of operated patients who underwent preoperative biopsy has not significantly changed over the study period (2008-14 vs. 2015-17: 67% vs. 60%; p = 0.779). Of the operated patients, 38 (69%) were resected, 15 (27%) were found to have unresectable PHC and 2 patients (4%) were found to have benign disease intraoperatively, which was subsequently confirmed by clinical and radiological follow-up. 5 out of 19 patients (26%) who underwent resection without a positive preoperative biopsy had benign disease on postoperative histology (post-inflammatory stricture [N = 3], gallstone disease [N = 1], IgG4-related cholangiopathy [N = 1]). The proportion of resected patients who had benign disease has not significantly changed over time (2008-14 vs. 2015-7; 5% vs. 25%; p = 0.140). The rate of benign disease in operated and resected patients was 7/55 (13%) and 5/38 (13%), respectively.

Of the 100 patients who did not undergo surgery, 68 had biopsy confirmed PHC, but were either unfit for surgery or had inoperable disease on imaging. The remaining 32 patients underwent clinical and radiological follow-up after initial evaluation (including negative biopsy in 19 patients). During the follow-up period, 14 patients (44%) were subsequently diagnosed with cholangiocarcinoma based on clinical and/or radiological progression and underwent palliative treatment. The remaining 18 patients (56%) were considered to have benign disease based on their clinical course after a median follow-up of 44 months (range 12–77).

To summarise, of the entire cohort (N = 155), 25 patients (16%) had a final diagnosis of benign disease based on either resection histology (N = 5), or clinical/radiological follow-up (N = 20). Benign disease was present in 7/55 (13%) operated patients (including 5/38 resected patients) and 18/100 non-operated patients (18%) (Fig. 1).

Predictors of benign proximal biliary stricture

Univariable analysis of clinical factors (Table 1) found no significant associations between the final diagnosis (benign vs. malignant PBS) and age (p = 0.475), gender (p = 0.663), abdominal pain (p = 0.055) or clinical jaundice (p = 0.174). However, those patients with benign PBS were found to have significantly lower levels of biochemical markers at referral than those with malignant disease, with median serum bilirubin of 42 vs. 168 μmol/l (p = 0.012) and CA19.9 of 18 vs. 343 U/ml (p < 0.001). Of the radiological factors, no significant associations were detected between the final diagnosis and ductal separation (p = 0.805), lesion/duct enhancement (p = 0.126), or the number (p = 0.424) or length (p = 0.056) of strictures. However, benign PBS were significantly less likely to be associated with the presence of a mass on CT (36% vs. 69%; p = 0.003).

A multivariable analysis was then performed to identify independent predictors of benign PBS. CA19.9 levels were unavailable in 28% of patients at referral, who would have been excluded from the multivariable analysis, as this used a complete-cases approach. In light of this, and in order to maximise the included sample size, a "missing data" group was produced for analysis. Patients in whom CA19.9 was available were dichotomised using 100 U/ml as a cut-off value. The resulting model (Table 2) found low CA19.9 to be the strongest predictor of benign PBS, with an odds ratio of 10.35 (p = 0.001) for those with levels <100 vs. 100+ U/ml at referral. Patients with missing data for CA19.9 were found to have a similar benign PBS risk than those with CA19.9 < 100 U/ml. Abdominal pain (OR: 3.36, p = 0.027) or absence of mass on CT (OR: 4.66, p = 0.004) were both significant independent predictors of benign PBS, whilst the absence of lesion/duct enhancement narrowly missed significance in the model (OR: 2.96, p = 0.055).

Discussion

In this study, we have identified several independent risk factors that may predict benign disease in a consecutive series of 155 patients with proximal biliary strictures. The results from this study may allow risk stratification and assist decision making in potential surgical candidates. According to an international consensus meeting in 2014, preoperative biopsy is not required

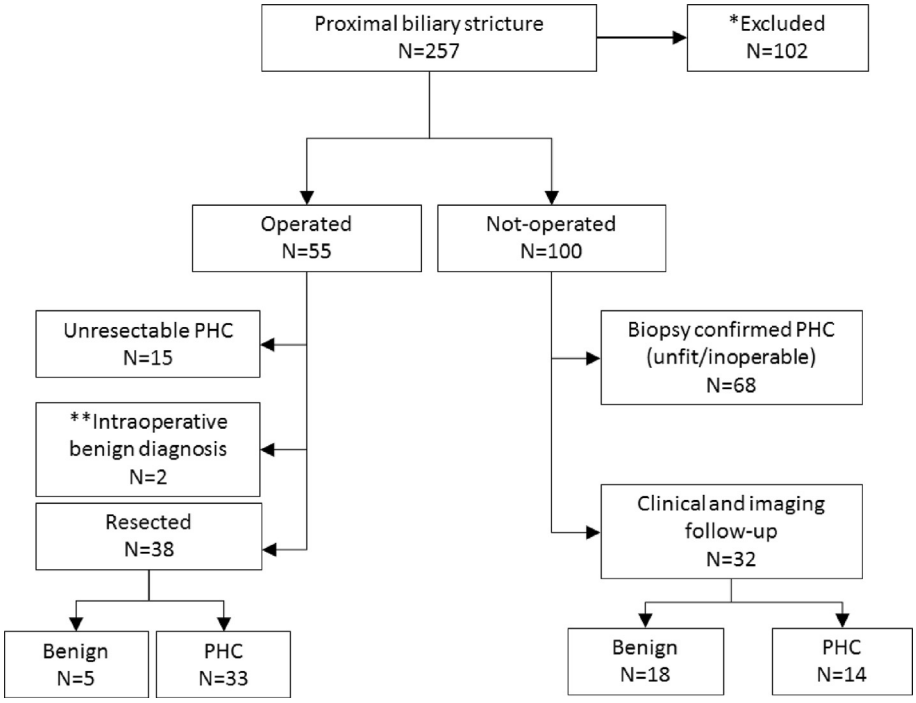

Figure 1 Flow chart summarising the outcome of patients referred with a proximal biliary stricture. PHC: Perihilar cholangiocarcinoma. *Extrinsic hilar obstruction (30), history of previous malignancy (2), metastatic PHC (27), primary sclerosing cholangitis (21), post-cholecystectomy stricture (1), intrahepatic or distal biliary stricture (11), common bile duct stone (1), intraductal papillary neoplasm (1), imaging not available (3) or no final diagnosis (5). ** The two patients with provisional diagnoses of benign proximal biliary stricture during surgery subsequently entered clinical and imaging follow-up where this diagnosis was confirmed

Table 1 Comparison of clinical and radiological factors according to final diagnosis

	N	Final diagnosis		p-value
		Benign	Malignant	
Clinical and biochemical factors				
Age (Years)	155	67 (57–73)	68 (61–73)	0.475
Gender (Male)	155	15 (60%)	70 (54%)	0.663
Abdominal pain	155	11 (44%)	32 (25%)	0.055
Clinical jaundice	155	17 (68%)	106 (82%)	0.174
Serum bilirubin (μmol/l)	152	42 (14–182)	168 (47–283)	**0.012**
CA 19.9 (U/ml)	112	18 (10–87)	343 (105–1179)	**<0.001**
Radiological factors				
Ductal separation	155	18 (72%)	97 (75%)	0.805
Number of strictures	155			0.424[b]
1		20 (80%)	113 (87%)	
2		4 (16%)	14 (11%)	
3		1 (4%)	3 (2%)	
Longest stricture (mm)	154[a]	19 (13–32)	26 (19–35)	0.056
Lesion/duct enhancement	155	8 (32%)	65 (50%)	0.126
Presence of mass on CT	155	9 (36%)	90 (69%)	**0.003**

Data are reported as N (%), with p-values from Fisher's exact tests, or as median (IQR), with p-values from Mann–Whitney tests, unless stated otherwise. Bilirubin and CA 19.9 were measured at referral. Bold p-values are significant at $p < 0.05$.
[a] Stricture length was unmeasurable in one patient, since a stent was in-situ.
[b] p-Value from a Mann–Whitney test, as the factor is ordinal.

prior to proceeding with attempted resection of suspected perihilar cholangiocarcinoma (PHC), "provided that benign aetiologies have been excluded".[6] However, there are currently no evidence-based guidelines that allow clinicians to differentiate benign and malignant strictures with a high degree of accuracy. The surgical management of PHC typically involves an extended hepatectomy to achieve negative margins, and is associated with a significant risk of morbidity (up to 43%) and mortality (up to 15%) even in experienced centres.[9,10] Up to 15% of patients with suspected PHC will be found to have benign pathology on postoperative histology,[2,4,8–10] and such patients are unnecessarily exposed to the risks of major hepatectomy. There is therefore an urgent need to stratify patients according to the risk of malignancy, and our analysis has identified an association between serum bilirubin, CA19.9 and the presence of a mass on CT scan with a malignant stricture.

Table 2 Multivariable analysis of predictors of benign stricture

	Odds ratio	p-value
Abdominal pain	3.36 (1.15–9.81)	**0.027**
CA19.9 at referral[a]		**0.004**
<100 U/ml	10.35 (2.53–42.42)	**0.001**
100+ U/ml	–	–
Missing data[a]	8.06 (1.94–33.56)	**0.004**
Absence of lesion/duct enhancement	2.96 (0.98–8.99)	0.055
Absence of mass	4.66 (1.65–13.13)	**0.004**

Results are from a multivariable binary logistic regression model. All factors from Table 1 were considered for inclusion, with a backwards stepwise procedure used for variable selection. Bold p-values are significant at $p < 0.05$.
[a] Since CA19.9 had a large quantity of missing data, cases where data were unavailable were grouped into a "missing data" category, to prevent them from being excluded.

Our unit's current policy is to perform a preoperative biopsy (percutaneous endobiliary or endoscopic ultrasound-guided) in all potential surgical candidates. However, given that the sensitivity of biopsy is only 69–75%,[7] a significant proportion of patients with PHC will have a negative biopsy, and a repeat biopsy may delay potentially curative surgery. Patients at high risk of PHC (absence of abdominal pain, high CA19.9 and mass on CT) should be considered for surgical exploration even after a negative biopsy. Conversely, patients at low risk of PHC (abdominal pain, low CA19.9 and no mass on CT) are likely to have benign disease and should undergo at least one repeat biopsy in addition to interval imaging and clinical follow-up.

Our study found that the tumour marker CA19.9 was an independent predictor of malignancy in PBS. The association between raised CA19.9 and pancreatobiliary cancers is well established, although when used in isolation it has a low sensitivity as a diagnostic test.[14–16] In a study of 58 patients, Saluja *et al.* also reported a cut-off value of CA19.9 $\geq$100.[16] A study by Charatcharoenwitthaya *et al.* also demonstrated a significant association between CA19.9 and malignancy in patients with PSC. Our study excluded patients with PSC, as it was felt that they represented a unique patient group with a distinct diagnostic/treatment pathway compared to patients presenting with a *de novo* proximal biliary stricture. The results of our study are therefore not applicable to PSC patients, and the predictors of malignancy in this patient cohort warrants further study.

Several studies have reported a significant association between serum bilirubin and malignancy,[11–13] and this is presumably because malignant strictures are generally narrower than benign strictures. However, whilst serum bilirubin was found to be significantly predictive of benign disease on univariable analysis, this association was weaker than that of CA19.9. Since these two markers were also significantly correlated (Spearman's rho: 0.41, $p < 0.001$), only CA19.9 was identified as an independent predictor of benign disease on multivariable analysis.

Cross-sectional imaging is an essential component of the work-up of patients with suspected PHC, and provides both diagnostic and staging information. Consensus guidelines recommend performing either high resolution CT with arterial and portal venous imaging or MRI/MRCP.[6] In our study, prospective assessment of imaging was performed by an experienced

HPB radiologist who was blinded to the final diagnosis, and we found that the presence of either a mass or duct/lesion enhancement were both independent predictors of malignancy. The fact that imaging was reviewed by a blinded radiologist suggests that these diagnostic features are objective and reliable predictors of malignancy, but it is unclear whether these findings would be reproducible in a non-specialist centre. Further work would be needed to investigate for inter-observer variability between specialist and non-specialist radiologists, but it must be emphasised that all patients presenting to a non-specialist hospital with a PBS should undergo cross-sectional imaging *before* biliary stenting,[6] and the imaging should be reviewed by an experienced HPB radiologist at a specialist MDT.

This study has several limitations. Although the sample size in our study is large compared to other single centre studies,[16,17] the number of patients in our cohort with benign disease is small, due to the rarity of perihilar cholangiocarcinoma. As such, the statistical power of the analyses performed will be suboptimal, potentially leading to an increased false-negative rate. Despite this, multivariable analyses identified multiple significant independent predictors of benign disease, implying that statistical power was sufficient to identify at least some of the predictors of this outcome. However, there may be other, weaker predictors of benign disease that were not detected by the analysis. The small sample size may also have resulted in a degree of overfitting in the multivariable analysis, especially since the "one in ten" rule was exceeded by including four covariates, but having only N = 25 events (i.e. benign strictures). As such, the reliability of the multivariable model and, hence, its generalizability to other cohorts cannot be confirmed. In light of the increased risks of both false-negatives and overfitting in the multivariable analysis, we would recommend that the findings should be validated in a larger cohort, which could potentially be recruited by multicentre international collaboration between high volume specialist centres.

Although our study identified all consecutive patients with PBS who had been referred to a specialist MDT, due to its retrospective design there was some missing data, particularly for serum CA19.9, which was only available in 72% of patients. This is a drawback of the study, especially if selection bias was present in the decision to measure this variable. In an attempt to mitigate this effect, those with missing data were treated as a separate category of CA19.9 in the multivariable model, to prevent cases from being excluded from the model, which suggested that the rates of benign disease in this group were consistent with those with levels <100 U/ml.

Despite the limitations detailed above, the authors believe that the risk factors that were identified provide practical guidance to clinicians involved in the management of cholangiocarcinoma, and give a basis for further work to develop and validate a risk score in a larger cohort.

In conclusion, this study has identified several independent variables that may differentiate benign and malignant proximal biliary strictures. A larger multi-institutional study is needed to confirm these findings, and to develop a risk score to stratify patients with suspected PHC.

Conflict of interest

None declared.

References

1. Vauthey J, Blumgart LH. (1994) Recent advances in the management of cholangiocarcinomas. *Semin Liver Dis* 14:109–114.
2. Corvera CU, Blumgart LH, Darvishian F, Klimstra DS, Dematteo R, Fong Y *et al*. (2005) Clinical and pathologic features of proximal biliary strictures masquerading as hilar cholangiocarcinoma. *J Am Coll Surg* 201:862–869.
3. Nakazawa T. (2019) Difficulty in the diagnosis of isolated IgG4-related sclerosing cholangitis. *Dig Endosc* 31:391–392.
4. Uhlmann D, Wiedmann M, Schmidt F, Kluge R, Tannapfel A, Berr F *et al*. (2006) Management and outcome in patients with Klatskin-mimicking lesions of the biliary tree. *J Gastrointest Surg* 10:1144–1150.
5. Khan SA, Davidson BR, Goldin RD, Heaton N, Karani J, Pereira SP *et al*. (2012) Guidelines for the diagnosis and treatment of cholangiocarcinoma: an update. *Gut* 61:1657–1669.
6. Mansour JC, Aloia TA, Crane CH, Heimbach JK, Nagino M, Vauthey J. (2015) Hilar cholangiocarcinoma: expert consensus statement. *HPB* 17: 691–699.
7. Mokham K, Malik Y, Derosas C, Isaac J, Marudanayagam R, Mehrzad H *et al*. (2017) Percutaneous transhepatic cholangiographic endobiliary forceps biopsy versus endoscopic ultrasound fine needle aspiration for proximal biliary strictures: a single-centre experience. *HPB* 19:530–537.
8. Gerhards MF, Vos P, Van Gulik TM, Rauws EAJ, Bosma A, Gouma DJ. (2001) Incidence of benign lesions in patients resected for suspicious hilar obstruction. *Br J Surg* 88:48–51.
9. Kloek JJ, van Deldein OM, Erdogan D, ten Kate FJ, Rauws EA, Busch OR *et al*. (2008) Differentiation of malignant and benign proximal bile duct strictures: the diagnostic dilemma. *World J Gastroenterol* 14: 5032–5038.
10. Roos E, Hubers LM, Coelen RJS, Doorenspleet ME, De Vries N, Verheij J *et al*. (2018) IgG4-associated cholangitis in patients resected for presumed perihilar cholangiocarcinoma: a 30-year tertiary care experience. *Am J Gastroenterol* 113:765–772.
11. Bain V, Abraham N, Jhangri G, Alexander T, Henning R, Hoskinson M *et al*. (2000) Prospective study of biliary strictures to determine the predictors of malignancy. *Can J Gastroenterol* 14:397–402.
12. Al-Mofleh IA, Aljebreen AM, Al-Amri SM, Al-Rashed RS, Al-Faleh FZ, Al-Freihi HM *et al*. (2004) Biochemical and radiological predictors of malignant biliary strictures. *World J Gastroenterol* 10:1504–1507.
13. Garcea G, Ngu W, Neal CP, Dennison AR, Berry DP. (2011) Bilirubin levels predict malignancy in patients with obstructive jaundice. *HPB* 13: 426–430.
14. Are C, Gonen M, D'Angelica M, DeMatteo RP, Fong Y, Blumgart LH *et al*. (2006) Differential diagnosis of proximal biliary obstruction. *Surgery* 140:756–763.
15. Charatcharoenwitthaya P, Enders FB, Halling KC, Lindor KD. (2008) Utility of serum tumor markers, imaging, and biliary cytology for detecting cholangiocarcinoma in primary sclerosing cholangitis. *Hepatology* 48:1106–1117.

16. Saluja SS, Sharma R, Pal S, Sahni P, Chattopadhyay TK. (2007) Differentiation between benign and malignant hilar obstructions using laboratory and radiological investigations: a prospective study. *HPB* 9: 373–382.

17. Juntermanns B, Kaiser GM, Reis H, Saner FH, Radunz SVS, Vernadakis S *et al*. (2011) Klatskin-mimicking lesions: still a diagnostical and therapeutical dilemma? *Hepatogastroenterology* 58(106):265–269.