Diagnosis of bile duct cancer by bile cytology: usefulness of post-brushing biliary lavage fluid

Authors
Shinya Sugimoto1, Hiroyuki Matsubayashi1, Hirokazu Kimura2, Keiko Sasaki2, Kaori Nagata3, Sachiyo Ohno4, Katsushiko Uesaka5, Keita Mori6, Kenichiro Imai7, Kinichi Hotta7, Kohei Takizawa7, Naomi Kakushima7, Masaki Tanaka8, Noboru Kawata9, Hiroyuki Ono1

Institutions
Institutions are listed at the end of the article.

Background: Pathologic evidence of biliary diseases can be obtained from cytology in addition to endoscopic retrograde cholangiopancreatography (ERCP); however, the diagnostic effectiveness is not satisfactory.

Study aim: This retrospective, single-center study evaluated the efficacy of various sampling methods for the cytologic diagnosis of bile duct cancer.

Patients and methods: Biliary samples included bile that was simply aspirated, brush smear, brush-rinsed saline, and post-brushing biliary lavage fluid. A set of samples was compared for cytologic efficacy in 76 patients with surgically proven bile duct cancer and in 50 patients with benign biliary stricture.

Results: The cytologic sensitivity for diagnosing biliary cancer was 34% with aspirated bile, 32% with brush smear, 43% with brush-rinsed saline, and 70% with post-brushing biliary lavage fluid, in contrast to the null false-positive result in the benign cases. The sensitivity of cytology was significantly higher with post-brushing lavage fluid than with the other three sampling methods ($P<0.0001$), and post-brushing lavage fluid improved the cumulative sensitivity by 24% ($P=0.002$). The sensitivity of biliary cytology was also associated with the amount of aspirated bile ($P=0.01$) and with the aspiration site ($P=0.03$). The rate of cancer positivity in a cytology set differed according to the tumor macroscopic type (85% in the protruding type vs. 40% in the flat type; $P=0.003$), and according to the size of the cancer (87% for tumors $\geq 50$ mm vs. 66% for tumors <50 mm; $P=0.02$).

Conclusions: Post-brushing biliary lavage fluid cytology provides superior diagnostic efficacy, and its addition to ERCP procedures is recommended for obtaining cytologic evidence of bile duct cancer.

Introduction

Several disorders are involved in the etiology of bile duct strictures, and their differential diagnosis is often difficult when based only on clinical images. Pancreatic head cancer and bile duct cancer account for the most cases of malignant biliary strictures. The histologic diagnosis of pancreatic head cancer is often confirmed with endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA), whereas pathologic evidence for bile duct cancer is usually obtained from standard endoscopic trans-papillary tissue samplings, such as cytology and forceps biopsy. However, an increase in the number of cases of benign biliary stricture masquerading as biliary carcinoma has recently been indicated [1 – 4]. Many of these studies report that even after diagnostic attempts with cytopathologic examinations, surgical resection is still performed. Therefore, better accuracy is needed for biliary cytology.

Endoscopic retrograde cholangiopancreatography (ERCP) conducted for biliary drainage can easily incorporate biliary cytology, but only a few sampling methods have been reported [5 – 7] for a small number of cases. The use of a brush can improve the diagnostic ability of biliary cytology by providing samples such as brush smears and brush-rinsed saline. However, even with this type of sample, the sensitivity for the diagnosis of bile duct cancer is generally unsatisfactory (23% – 63%) [8 – 12]. In addition, most of the previous studies were conducted without the availability of surgical samples, so the cytologic analysis was not compared with the pathologic findings in the resected biliary cancers.

In the present study, we analyzed the biliary cytology of a large number of bile duct cancers and compared the efficacy of biliary cytology for four types of biliary sampling methods: simple aspiration of bile, two stricture-scraping brush methods (brush smear and brush-rinsed saline), and a no-
vel sampling of post-brushing biliary lavage fluid. The cytologic efficacy was also compared with the sampling and pathologic factors determined in the resected materials from the same cases.

**Patients and methods**

This was a single-center retrospective study. We retrospectively investigated the diagnostic sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy of the biliary cytology obtained with each sampling method. Demographic factors, sampling-related factors (amount of aspirate and site of aspiration), and tumor pathologic factors were analyzed for their association with the efficacy of biliary cytology.

**Patients**

The present study retrospectively investigated 76 patients with bile duct cancer who underwent all four types of bile duct sampling methods and surgery at the Shizuoka Cancer Center between January 2008 and July 2013 (Table 1). The study also included 50 patients with benign biliary stricture as a control group (33 with autoimmune pancreatitis-associated sclerosing cholangitis [AIP-SC], 11 with benign biliary stricture with choledocholithiasis, 4 with primary sclerosing cholangitis or immunoglobulin G4 (IgG4)-related sclerosing cholangitis, 2 with benign biliary anastomosis at a hepaticojejunostomy). All cases of bile duct cancer were diagnosed pathologically by mapping thin-sliced resected tissues. The tip of the brush was then rubbed onto a slide glass, and the sample on the slide was immediately fixed in 100% ethanol (brush smear). The brush was thoroughly washed in 10 mL of physiologic saline, and the saline was used as a third sample (brush-washed saline or brush-rinsed saline). Finally, after the brushing, the biliary lumen was washed with approximately 5 to 15 mL of physiologic saline, which was then re-collected via the outer sheath of the brush (post-brushing biliary lavage fluid). In patients with symptoms of cholangitis (fever, abdominal pain, high white blood cell count, tachycardia, decreased blood pressure), we washed the biliary lumen with 5 mL of physiologic saline. Otherwise, we washed with approximately 15 mL of physiologic saline. Finally, the aspirated bile, brush smear, brush-rinsed saline, and post-brushing biliary lavage fluid samples were examined cytologically for 76 patients with bile duct cancer and for 49, 25, 25, and 17 patients, respectively, with benign bile duct stricture.

**Grading of biliary cytology**

All specimens were stained with Papanicolaou solution and rated as class I to V. Cytologic findings considered suggestive of malignancy were the following: irregularly arranged nuclei, irregular margins for clusters, enlarged nuclei, irregularly shaped nuclei, and abnormal chromatoin [15, 16]. Other notable findings were a necrotic background and cell clusters of different sizes. The biliary cytology was graded according to the combinations of these cytologic findings. The diagnostic grading system used in the present study was as follows: inadequate; benign (class I) (Fig. 2a); atypical cells present (class II); suspicious for cells from a neoplastic lesion (class III) (Fig. 2b); strongly suspicious for malignancy (class IV); and definitely malignant (class V) (Fig. 2c). In this study, classes I through III were regarded as benign, and classes IV and V as malignant (positive for cancer). Cytologic evaluation was performed before the surgical resection in all cases by a single specialist (K. S., listed as a co-author) in the field of hepatobiliary-pancreatic pathology.

**Pathologic diagnosis of the surgical specimen**

We analyzed the following pathologic factors affecting the efficacy of biliary cytology in 76 patients with surgically resected biliary cancer: location, size, macroscopic type, differentiation, depth of invasion, lymph duct permeation, venous permeation, lymph node metastasis, and invasion of adjacent organs (liver, gallbladder, and pancreas) [17]. The classification of the Japanese Society of Biliary Surgery [17] was used to categorize the macroscopic appearance of the bile duct cancers as papillary type, nodular type, or flat type. Cancers of the papillary type and nodular

---

**Table 1** Characteristics of the patients enrolled in a study of the sensitivity of bile cytology in the diagnosis of bile duct cancer.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Patients, n</th>
<th>Male patients, n</th>
<th>Median age (range), y</th>
<th>Median follow-up (range), mo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bile duct cancer</td>
<td>76</td>
<td>62</td>
<td>70 (48–83)</td>
<td>-</td>
</tr>
<tr>
<td>Hilar bile duct cancer</td>
<td>41</td>
<td>34</td>
<td>70 (56–83)</td>
<td>-</td>
</tr>
<tr>
<td>Common bile duct cancer</td>
<td>35</td>
<td>28</td>
<td>71 (48–80)</td>
<td>-</td>
</tr>
<tr>
<td>Benign bile duct stricture</td>
<td>50</td>
<td>42</td>
<td>67 (45–84)</td>
<td>40 (14–86)</td>
</tr>
<tr>
<td>AIP-SC</td>
<td>33</td>
<td>28</td>
<td>65 (45–84)</td>
<td>43 (14–77)</td>
</tr>
<tr>
<td>Choledocholithiasis</td>
<td>11</td>
<td>9</td>
<td>67 (61–81)</td>
<td>36 (20–86)</td>
</tr>
<tr>
<td>Sclerosing cholangitis</td>
<td>4</td>
<td>3</td>
<td>75 (71–82)</td>
<td>19 (14–59)</td>
</tr>
<tr>
<td>Stricture of hepaticojejunostomy</td>
<td>2</td>
<td>2</td>
<td>72 (66–78)</td>
<td>54 (51–57)</td>
</tr>
</tbody>
</table>

AIP-SC, sclerosing cholangitis associated with autoimmune pancreatitis.
type were grouped as protruding type. Tumor size, measured pathologically, included areas showing intramucosal (intraepithelial) spread of carcinoma, not just invasive cancer.

Statistical analyses
Factors possibly affecting cytologic sensitivity were first analyzed by univariate analysis with Fisher’s exact test and Wilcoxon’s rank sum test. The candidate factors were then re-analyzed by multivariate analysis with logistic regression analysis. All data analysis was performed with R Statistical Software, and a P value of less than 0.05 was considered statistically significant.

Results

Biliary cytology by each sampling method
Biliary cytology was analyzed by each sampling method in 76 patients with bile duct cancer and in 50 patients with benign biliary cytology (Table 2). The sensitivity of detecting bile duct cancer was 34% (26/76) with aspirated bile, 32% (24/76) with brush smear, 43% (33/76) with brush-rinsed saline, and 70% (53/76) with post-brushing biliary lavage fluid. A significant difference was noted between the sensitivity of cytology with post-brushing biliary lavage fluid and the other sampling methods (P<0.002) (Table 2). None of the samples obtained from the benign biliary strictures showed malignant cells, regardless of the sample type (100% specificity). The accuracy of cytology was best with post-brushing biliary lavage fluid (75%), and it was significantly higher than the accuracy obtained with aspirated bile (60%; P=0.02), brush-rinsed saline (57%; P=0.01), or brush smears (49%; P=0.0002) (Table 2).

Technical factors affecting the efficacy of biliary cytology
We considered that some of the technical aspects of biliary sampling, such as amount of aspirate and site of aspiration, might affect the efficacy of biliary cytology (Table 3). Data on aspirate amount were available for 75 cases of aspirated bile and for 67 cases of post-brushing biliary lavage fluid, and data on the aspiration site were available for 75 cases of aspirated bile and for 75 cases of post-brushing biliary lavage fluid. Cases without the records for these data were excluded from this analysis. The receiver operating characteristic curve analysis indicated that the required amount of aspirate was 10mL for aspirated bile and 20mL for post-brushing biliary lavage fluid. The sensitivity of detecting cancer cells in 10mL or more of aspirated bile (52%, 12 of 23 cases) was significantly higher than the sensitivity in less than 10mL (27%, 14 of 52 cases; P=0.03) (Table 3). The sensitivity of post-brushing biliary lavage fluid did not significantly differ according to the aspirate amount (80% in ≥20mL vs. 65% in
Table 2  Sensitivity of biliary cytology by sampling step (n = 76).1

<table>
<thead>
<tr>
<th>Sample</th>
<th>Median amount of specimen (range), mL</th>
<th>Sensitivity, % (n)</th>
<th>Cumulative sensitivity, % (n)</th>
<th>PPV, % (n)</th>
<th>NPV, % (n)</th>
<th>Accuracy, % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirated bile</td>
<td>5 (1-40)</td>
<td>344 (26/76)</td>
<td>34 (26/76)</td>
<td>100 (26/26)</td>
<td>49 (49/99)</td>
<td>605 (75/125)</td>
</tr>
<tr>
<td>Brush smear</td>
<td></td>
<td>325 (24/76)</td>
<td>45 (34/76)</td>
<td>100 (24/24)</td>
<td>32 (25/77)</td>
<td>496 (49/101)</td>
</tr>
<tr>
<td>Brush-rinsed saline</td>
<td></td>
<td>43 (33/76)</td>
<td>55 (42/76)</td>
<td>100 (33/33)</td>
<td>37 (25/68)</td>
<td>577 (58/101)</td>
</tr>
<tr>
<td>Post-brushing biliary lavage fluid</td>
<td>7.5 (2-40)</td>
<td>70 (53/76)</td>
<td>79 (60/76)</td>
<td>100 (53/53)</td>
<td>43 (17/40)</td>
<td>75 (70/93)</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value.
1 This analysis was done in 76 patients who underwent all four types of biliary sampling.
2 a vs. d, b vs. d, e vs. h, f vs. h: P < 0.0001.
3 i vs. l, P = 0.02.
4 j vs. l, P = 0.0002.
5 c vs. d: P = 0.001.
6 g vs. h: P = 0.002.
7 k vs. l, P = 0.01.

Adverse events
No obvious adverse events, such as biliary hemorrhage and perforation, were noted for the brushing procedures in any of the cases.

Take-home message
The aim of this study was to reveal the usefulness of post-brushing biliary lavage fluid cytology in the diagnosis of cholangiocarcinoma. The biliary lavage method demonstrated a significantly higher sensitivity (70%) than standard biliary cytology with aspirated bile (34%) and/or brushing samples (32% and 43%). The amount of aspirate, tumor size (including intramucosal spread), and macroscopic tumor type all affected the efficacy of biliary cytology.

Discussion
This study demonstrated a high sensitivity (79%) and specificity (100%) for the cumulative results obtained with four cytologic techniques. Post-brushing biliary lavage fluid cytology added 24% (18/76) to the sensitivity compared with the cumulative results obtained with the other three cytologic techniques (P = 0.002) (Table 2). This sensitivity was higher than previously reported values (23%–63%) [8–12]. However, simple comparison of sensitivities among different institutions does not provide an objective evaluation of a methodology because of the wide range of interobserver variation that has been reported in the cytologic and pathologic diagnosis of disease in many organs [16, 18–21]. This diagnostic disagreement is associated with the experience [18] and institution type [19] of the pathologists and with the pathologic type of tumor [20]. Discordance in cytologic criteria is seen even among different nations [22, 23].
This is also true in the field of biliary cytology for the diagnosis of bile duct cancer, in which a rate of interobserver agreement of 48% to 52% was reported in a previous study (κ=0.66) [21]. We therefore performed a more objective analysis by sampling the bile of 76 patients with biliary cancer and comparing the diagnostic efficacy of four types of biliary sampling. Post-brushing biliary lavage fluid cytology (70%) demonstrated more than twice the sensitivity of aspirated bile cytology (34%) and added a further 24% to the cumulative sensitivity of the other three sampling methods (Table 2).

Several methodologic devices have been reported for the improvement of biliary cytology based on the aspects of sampling, cytopreparation [24], and molecular approach [25–27]. A brush [15,28] is the most common sampling device used in the daily clinic, and a more cellular sample is usually obtained with brushing than with direct bile sampling as a consequence of traumatization of the ductal epithelium through contact with the abrasive bristles of the brush. Although the specificity is high, the sensitivity for the detection of malignancy remains unsatisfactory (23%–63%) [8–12,15,28]. The use of a larger cytology brush (3mm×5cm) increased the cellularity but did not improve the cancer detection rate over that obtained with a standard brush [9]. The combination of biliary dilation, endoscopic needle aspiration, and subsequent brushing cytology showed a higher sensitivity (85%) than brushing alone (57%; P<0.02) [29].

The use of a liquid-based cytopreparation, or ThinPrep [15, 16], improved interobserver agreement regarding cytologic features for diagnosing malignancy. In molecular diagnostic approaches, reported by Krishnamurthy et al. [25], DNA histograms decreased the false-negative rate of biliary cytology by 7%, while DNA hyperploidy (DNA content >5c) diagnosed bile duct cancer with 62% sensitivity and 91% specificity. Kipp et al. [26] used fluorescence in situ hybridization (FISH) on chromosomes 3, 7, 9, and 17 to evaluate biliary brushing samples obtained from biliary strictures and reported 34% sensitivity for detecting cancer by FISH and 15% by cytology (P<0.01). In the current study, the evaluation of post-brushing biliary lavage fluid, which can be collected in 1 minute by washing the biliary lumen with saline, demonstrated a significant improvement in biliary cytology without adverse events. These other molecular methods, if combined with the use of post-brushing biliary lavage fluid, may result in even greater improvements in sensitivity to an excellent level.

In the current study, the size and macroscopic appearance of the cancer were strongly correlated with the efficacy of biliary cytology (Table 4). The theoretical mechanism to explain the effectiveness of the biliary lavage method is that the coupling of exfoliation of the epithelial cells by brush stimulation and lavage of the biliary surface with saline increases the chance of collecting malignant cells via the catheter (Table 2). If this is the case, then the amount of superficial spread of intramucosal carcinoma is important for this cytology, but the invasiveness of the cancer is not. The pathologic observation of 253 resected cases showed that the mucosal extension of carcinoma over a macroscopic mass was observed in 47% to 48% of cases of biliary cancer, with a median length of 10mm from the invasive margin [30]. The characteristics of biliary cancer with widely spreading intramucosal carcinoma were macroscopically protruding type, less advanced stage, slower growth, and better prognosis [31]. Conversely, flat invasive cancers with minimal mucosal spreading, which are associated with aggressive biological behavior, are difficult to detect with biliary cytology (Table 4).

Clinicopathologic studies of benign biliary strictures masquerading as biliary carcinoma are being increasingly reported [1–4]. Even with cytopathologic examination or demonstration of the known features of benign disease, resection continues to be performed because of the suspicion of a hidden biliary cancer. Avoiding these unnecessary surgeries and ensuring the timing of curable treatment will require improvements in biliary cytology in future studies.

Competing interests: None

Institutions
1. Division of Endoscopy, Shizuoka Cancer Center, Shizuoka, Japan
2. Division of Molecular Biology and Biochemistry, Graduate School of Medicine, Osaka University, Osaka, Japan
3. Division of Pathology, Shizuoka Cancer Center, Shizuoka, Japan
4. Division of Hepato-Biliary-Pancreatic Surgery, Shizuoka Cancer Center, Shizuoka, Japan
5. Division of Clinical Trials, Shizuoka Cancer Center, Shizuoka, Japan

Table 4 Factors affecting the cumulative sensitivity of biliary cytology (n = 76).1

<table>
<thead>
<tr>
<th>Factor</th>
<th>Condition</th>
<th>Sensitivity, % (n)</th>
<th>OR</th>
<th>95%CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Univariate</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td>≥ 70</td>
<td>79 (34/43)</td>
<td>1</td>
<td>0.35 – 3.01</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&lt;70</td>
<td>79 (26/33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>77 (48/62)</td>
<td>0.6</td>
<td>0.13 – 2.60</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>86 (12/14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macroscopic type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protruding</td>
<td></td>
<td>85 (56/66)</td>
<td>8.4</td>
<td>2.12 – 33.22</td>
<td>0.004</td>
</tr>
<tr>
<td>Flat</td>
<td></td>
<td>40 (4/10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size, mm</td>
<td>≥ 50</td>
<td>87 (41/47)</td>
<td>3.6</td>
<td>1.17 – 10.99</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>&lt;50</td>
<td>66 (19/29)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>Hepatic hiliar</td>
<td>88 (36/41)</td>
<td>3.3</td>
<td>1.05 – 10.27</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>Extrahepatic</td>
<td>69 (24/35)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth of invasion</td>
<td>m, fm ss, se</td>
<td>83 (5/6)</td>
<td>1.4</td>
<td>0.19 – 9.32</td>
<td>1</td>
</tr>
<tr>
<td>Diffentiation</td>
<td>Good, moderate</td>
<td>79 (3/4)</td>
<td>0.8</td>
<td>0.10 – 5.83</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>75 (57/72)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval; m, mucosal; fm, fibromuscular; ss, subserosal; se, exposed to serosa.

1 This analysis was done in 76 patients who underwent all four types of biliary sampling.
References

5 Farrell RJ, Jain AK, Brandwein SL et al. The combination of structure dilatation, endoscopic needle aspiration, and biliary brushings significantly improves diagnostic yield from malignant bile duct strictures. Gastrointest Endosc 2001; 54: 587 – 594
29 Farrell RJ, Jain AK, Brandwein SL et al. The combination of structure dilatation, endoscopic needle aspiration, and biliary brushings significantly improves diagnostic yield from malignant bile duct strictures. Gastrointest Endosc 2001; 54: 587 – 594