Characterization of the urinary metabolic profile of cholangiocarcinoma in a United Kingdom population

Munirah Alsaleh1
Thomas A Barbera1
Helen L Reeves2
Matthew E Cramp3
Stephen Ryder4,5
Hani Gabra1,6
Kathryn Nash7
Yi-Liang Shen1,8
Elaine Holmes1
Roger Williams9
Simon D Taylor-Robinson1

1Division of Surgery and Cancer, Imperial College London, London; 2Northern Institute for Cancer Research, Medical School, University of Newcastle, Newcastle upon Tyne, UK; 3Liver Unit, Derriford Hospital, Plymouth, Devon, UK; 4Nottingham Digestive Diseases Centre, University of Nottingham, Nottingham, UK; 5NIHR Biomedical Research Unit, Nottingham University Hospitals NHS Trust, Queen’s Medical Centre, Nottingham, UK; 6Early Clinical Development, IMED Biotech Unit, AstraZeneca, Cambridge, UK; 7Liver Unit, Southampton General Hospital, Southampton, Hampshire, UK; 8Department of Radiation Oncology, Chang Gung Memorial Hospital and Chang Gung University, Taoyuan, Taiwan; 9Institute of Hepatology, UK

Correspondence: Simon D Taylor-Robinson
Liver Unit, St. Mary’s Hospital, London W2 1NY, UK
Tel +44 203 312 6199
Fax +44 207 924 9369
Email s.taylor-robinson@imperial.ac.uk

Background: Outside South-East Asia, most cases of cholangiocarcinoma (CCA) have an obscure etiology. There is often diagnostic uncertainty. Metabolomics using ultraperformance liquid chromatography mass spectrometry (UPLC-MS) offers the portent to distinguish disease-specific metabolic signatures. We aimed to define such a urinary metabolic signature in a patient cohort with sporadic CCA and investigate whether there were characteristic differences from those in patients with hepatocellular carcinoma (HCC), metastatic secondary liver cancer, pancreatic cancer and ovarian cancer (OCA).

Methods: Spot urine specimens were obtained from 211 subjects in seven participating centers across the UK. Samples were collected from healthy controls and from patients with benign hepatic disease (gallstone, biliary strictures, sphincter of Oddi dysfunction and viral hepatitis) and patients with malignant conditions (HCC, pancreatic cancer, OCA and metastatic cancer in the liver). The spectral metabolite profiles were generated using a UPLC-MS detector and data were analyzed using multivariate and univariate statistical analyses.

Results: The greatest class differences were seen between the metabolic profiles of disease-free controls compared to individuals with CCA with altered acylcarnitine, bile acid and purine levels. Individuals with benign strictures showed comparable urine profiles to patients with malignant bile duct lesions. The metabolic signatures of patients with bile duct tumors were distinguishable from patients with hepatocellular and ovarian tumors, but no difference was observed between CCA cases and patients with pancreatic cancer or hepatic secondary metastases.

Conclusion: CCA causes subtle but detectable changes in the urine metabolic profiles. The findings point toward potential applications of metabonomics in early tumor detection. However, it is key to utilize both global and targeted metabolomics in a larger cohort for in-depth characterization of the urinary metabolome in hepato-pancreato-biliary disease.

Keywords: cholangiocarcinoma, metabolomics, diagnostic biomarkers

Introduction
Cholangiocarcinoma (CCA) is a devastating malignancy. Tumors commonly present late in the course of the disease and with the exception of liver fluke-associated CCA in South-East Asia are sporadic in most cases.1 The commonest known risk factor of CCA in developed countries is primary sclerosing cholangitis (PSC).1 In England and Wales, the incidence of CCA has increased year on year.2

Given the diagnostic uncertainty that surrounds most CCA cases, a metabonomic approach might shed light not only on disease pathogenesis in CCA, but also
potentially provide information for diagnostic, predictive or prognostic use. Finding prognostic and diagnostic markers for CCA is urgent for liver fluke endemic regions, particularly for Thailand, where CCA is a major health problem, but it is also important for sporadic CCA, where the disease is fatal in majority of the cases. Due to the rarity of CCA in Western countries, it is difficult to acquire a sufficient number of samples to robustly identify and validate metabonomics-based molecular markers using large-scale studies. We have thus collected samples across the UK to build up numbers in partnership with the UK Clinical Research Network.

The primary objective of the current study was to evaluate differences between the urinary metabolite profiles of CCA cases, compared to healthy participants, patients with benign hepatic disease (gallstones, strictures, sphincter of Oddi dysfunction [SOD] and viral hepatitis) and patients with malignant conditions (hepatocellular, pancreatic and metastatic cancers). Ovarian cancer (OCA) urine samples were included as nonhepatic malignancy controls, as these tumors are of a completely different cell line and we hypothesized that they would have a very distinct metabolic profile from those of a hepatobiliary or pancreatic origin.

**Methodology**

**Patient and healthy volunteer recruitment**

Urine samples were collected from seven participating UK clinical liver centers in London, Manchester, Newcastle, Nottingham, Plymouth and Southampton and transported frozen to the Hepatology Biobank at St. Mary’s Hospital, London, UK. Sequentially presenting potential participants were identified and recruited by their clinician from both the in-patient or out-patient populations.

Inclusion criteria:

- Established histological and/or imaging diagnosis of any hepatopancreato-biliary malignancy, including:
  - CCA;
  - hepatocellular carcinoma (HCC);
  - pancreatic carcinoma;
  - liver metastases (any primary source).

- Established histological diagnosis of OCA, included as a malignant, nonhepatobiliary control.

- Established diagnosis of any benign hepatopancreato-biliary condition, including:
  - gallstone disease;
  - PSC;
  - SOD;
  - pre-cirrhotic hepatitis C infection
- Matched control individuals with no hepato-pancreato-biliary condition.

Healthy volunteers were sought from among visitors to the hospital, staff and students. After participants provided written informed consent, they were assessed at baseline for demographic data, medical history, drug history and dietary history. Ethical approval was obtained from Imperial College London REC, London, UK (REC Reference 09/H0712/82). This complied with the precepts set out in the 1975 Declaration of Helsinki on Human Rights.

**Sample processing**

A fasting urine specimen of 20 mL was collected from each participant. Urine specimens were kept on ice or in +4°C fridge while awaiting processing. Samples were centrifuged at 4°C at 1,000 g for 10 mins and then 2 mL of urine aliquots was transferred into 6 Eppendorf tubes (Eppendorf Ltd, Stevenage, UK), of which one was used for the current study. Urine samples were transported on dry ice to the South Kensington campus, Imperial College London for MS analysis.

**Quality control**

Quality control (QC) samples were prepared by pooling 50 μL of each urine sample into a Falcon tube (Sigma-Aldrich, Dorset, UK). A 200 μL aliquot was then transferred to an analytical MS well plate to enable acquisition of a QC spectrum every 10 samples.

**Chromatographic conditions**

The sample spectra were acquired using an ACQUITY™ ultraperformance liquid chromatography mass spectrometry (UPLC) system (Waters Ltd. Elstree, UK), coupled to an LCT Premier™ mass spectrometer (Waters MS Technologies, Ltd., Manchester, UK). RP-UPLC-MS was performed with electrospray ionization (ESI) in both positive and negative modes. The conditions were optimized using the QC samples in terms of peak shape, reproducibility and retention time.

**Tandem mass spectrometry**

Tandem mass spectrometry (MS/MS) analysis was performed using a quadrupole time-of-flight (TOF) Premier™ instrument
Collision-induced dissociation (CID) experiments of the QC sample were performed for structural elucidation of detected ions in each ionization mode. This was conducted subsequent to the original profiling run to save time and limit analytical variations in retention time and performance that can occur when returning to the instrument for CID analysis.

Two complementary MS/MS acquisition modes were used to ensure sufficient MS/MS coverage of ions of interest, data-dependent acquisition (DDA) and acquisition with no precursor ion selection, or data-independent acquisition (MS<sup>6</sup>). The DDA experiment was set to switch automatically from the MS to MS/MS mode using data-dependent criteria. It triggered MS/MS on the most abundant ions in each MS scan and provided fragments, which were specifically attributed to the precursor ion. In MS<sup>6</sup> mode, eluting peaks were subjected to both high and low collision energies in the collision cell of the mass spectrometer, with no prior precursor ion selection.<sup>4</sup>

**Metabolite assignment verification**

The molecular mass, retention time and fragmentation spectrum of the discriminant features were compared against online spectral libraries such as HMDB (www.hmdb.ca)<sup>5</sup> and METLIN (https://metlin.scripps.edu).<sup>6</sup> Metabolites were classified as either:

1) identified compounds confirmed with an authentic standard; 2) putatively annotated compounds (such as those based upon fragmentation pattern and/or spectral similarity with spectral databases); 3) putatively identified to match a certain chemical class (such as those based on spectral similarity to known compounds of a chemical class); or 4) as unknown compounds.

**Preprocessing**

The raw LC-MS data files were converted to CDV format by MassLynx<sup>TM</sup> version 4.1 application manager (Waters Corporation, Milford, CT, USA) and then imported into R Project version 3.1.0 (The R Foundation for Statistical Computing, 2014) for preprocessing using XCMS package version 2.14. (Bioconductor). Computational scripts written in-house were applied to: 1) filter and identify peaks; 2) correct for retention time drift; 3) match peaks across samples; and 4) fill in missing peaks.

**Statistical analysis**

SIMCA-P+ version 13.0.2 (Umetrics, Umeå, Sweden) was used for multivariate statistical analysis of the processed data. Initial analysis was performed using unsupervised principal component analysis (PCA) to explore variation in the dataset and examine clustering patterns or trends in the dataset, based on metabolic profile similarities or differences. Following PCA, orthogonal projections to latent structures discriminant analysis (OPLS-DA) was performed to maximize separation between predefined sample classes in order to view discriminatory features. Feature selection was based on the variable importance in projection (VIP) coefficients, which allow the X variables to be classified according to their explanatory power of Y (class information). Features with high VIP value, >1, were found to be the most relevant for explaining Y class information. The top 30 features were selected and identified for each model.

Validating multivariate models is essential to avoid overfitting the data. The model statistics, R<sup>2</sup>X, Q<sup>2</sup>Y, permutation test and CV-ANOVA p-value, were used to evaluate the model’s robustness. Permutation testing (with 100 permutations) was calculated for every OPLS-DA model using SIMCA-P+ version 13.0.2 (Umetrics). Univariate significance tests were then performed on the selected features using ANOVA with post-hoc testing (Tukey’s HSD), which is designed to account for multiple comparisons.

**Correlation with hierarchical clustering order**

R Project version 3.1.0 (The R Foundation for Statistical Computing, 2014) using corrplot package version 0.77 (CRAN) was used to perform hierarchical cluster analysis of Spearman’s correlation coefficient matrix. The cluster analysis was used to investigate correlations among the identified biochemical components. The correlation matrix was represented as a heatmap with rows and columns ordered according to a hierarchical clustering analysis. Hierarchical clustering is an unsupervised method. No class information was given to calculate the model, which was found to be suitable for exploratory data analysis. Positively and negatively correlated analytes were displayed in blue and red colors, respectively. A circle was used to represent correlations between pairs of compounds. The circle diameter and color intensity were proportional to the correlation coefficients and indicate statistically significant correlations (<0.05). The circle diameter and color intensity were proportional to the correlation coefficients.
Results

Demographics, clinical data and cohort description

A total of 211 subjects provided urine samples for analysis using global LC-MS metabolic profiling. Table 1 shows the demographics and clinical description of the study cohort. Overall, the samples were categorized into 3 groups: healthy controls, participants with benign hepatic conditions (including SOD, strictures, gallstones and viral noncirrhotic hepatitis C [HCV]) and participants with malignant tumors (including CCA, HCC, pancreatic cancer, OCA and liver metastases cases). Of the 10 participants with CCA, 4 were diagnosed with perihilar CCA, 2 with distal CCA, 1 case with intrahepatic CCA and 3 cases with unknown tumor origin. Healthy controls (n=22) were younger than those with both benign and malignant hepatic conditions. With respect to gender, the HCC group comprised mostly of men (85.7%), followed by the pancreatic cancer patient group with 71.4% male participants. The participants’ racial background was diverse, but they were mostly white European.

CCA patients vs benign liver disease cases

The urinary spectral data from CCA patients were compared to benign liver disease cases using OPLS-DA analysis (Figure 1–5). The most discriminant OPLS-DA model was the one between the urinary MS spectral profile from CCA cases, compared to healthy participants (Figure 1). Similarly, profiles from patients with SOD and gallstones were distinguishable from CCA patients (Figures 2 and 3). The urinary spectral data from participants with benign strictures, including PSC cases, were poorly differentiated from CCA patients (Figure 4). The model comparing CCA patients versus a disease control group – those with pre-cirrhotic chronic hepatitis C infection – did show profile separation (Figure 5).

Altered urinary metabolites between CCA patients and benign cases

The urinary molecules influencing the OPLS-DA model separation between CCA cases and benign disease conditions (SOD, HCV and gallstones) were similar to those identified from the model comparing CCA to healthy participants. Therefore, to avoid repetition of data, Tables 2 and 3 list the metabolites driving the separation between the urinary metabolic profiles from CCA patients compared to healthy controls. Differences in these metabolic features between the groups were then evaluated using ANOVA with post hoc testing. The relevant metabolites are further illustrated in Figures 6–9.

Multigroup PLS-DA analysis was calculated to further assess the molecular panel underpinning the urinary metabolic differences between the six distinct conditions. The statistical model generated scores plot and a corresponding plot which rank metabolites, based on their VIP scores. Based on the crossvalidation statistics, differences in the urinary metabolic signature were identified in the ESI+ spectral data (Figure 10), but not the ESI− spectral data.

CCA patients vs other cancers

The MS data acquired using urine samples from CCA patients were compared to those acquired from patients with malignant pathologies. Patients with CCA showed a distinguishable urinary metabolic profile from patients with hepatocellular (Figure 11) and ovarian tumors (Figure 12), but no difference was observed between CCA cases and patients with pancreatic and metastatic cancers. The urine profiles from patients with pancreatic tumors also did not differ from metastatic cases.

Discussion

The study identified a coherent metabolic pattern associated with bile duct tumors. This molecular pattern was similar to the one associated with participants with pancreatic and metastatic cancers, yet distinct from individuals with hepatocellular and ovarian tumors. The urine metabolic profiles from CCA patients were also distinguishable from patients with benign liver pathologies (including SOD, gallstones and chronic noncirrhotic hepatitis C). However, the urinary metabolome from individuals with benign biliary strictures revealed a relatively similar underpinning metabolic process, compared to malignant strictures.

As expected, the best separation in the urinary metabolome was achieved between the spectral profiles from CCA patients and healthy controls. Interestingly, the model comparing the profiles of patients with CCA and benign strictures (including PSC cases) did not achieve statistical significance. Yet, participants with benign strictures showed a distinct molecular pattern from healthy controls.

Metabolites driving the separation between the two classes (healthy vs benign biliary strictures) were relative similarity similar to those separating CCA and healthy controls with depletion in urinary acylcarnitine and creatinine associated with patients with biliary strictures. In the ESI+, a decrease in the levels of urinary p-cresol sulfate, indoxylsulfate,
Table 1 Demographics of study population

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Benign</th>
<th>Malignant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Healthy</td>
<td>SOD</td>
</tr>
<tr>
<td>Participants, n</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>Age (range), years</td>
<td>34 (24–58)</td>
<td>55 (30–74)</td>
</tr>
<tr>
<td>Male, %</td>
<td>45.4</td>
<td>37.5</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>25 (19.5–39)</td>
<td>n/a</td>
</tr>
<tr>
<td>Serum biochemistry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilirubin, μmol/L</td>
<td>n/a</td>
<td>10.1 (4–21)</td>
</tr>
<tr>
<td>ALT, IU/L</td>
<td>n/a</td>
<td>53 (17–99)</td>
</tr>
<tr>
<td>ALP, IU/L</td>
<td>n/a</td>
<td>94 (20–184)</td>
</tr>
<tr>
<td>Albumin, g/L</td>
<td>n/a</td>
<td>27.6 (1–41)</td>
</tr>
<tr>
<td>Urea, mmol/L</td>
<td>n/a</td>
<td>26.5 (3.2–109)</td>
</tr>
<tr>
<td>Creatinine, μmol/L</td>
<td>n/a</td>
<td>66 (54–83)*</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Black</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Asian-Indian</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Not available</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Note: *Missing data.
Abbreviations: ALT, alanine transaminase; ALP, alkaline phosphatase; CCA, cholangiocarcinoma; GS, gallstones; HCC, hepatocellular carcinoma; HCV, noncirrhotic hepatitis C; METS, metastatic cancer of the liver; OCA, ovarian cancer; PCA, pancreatic cancer; SOD, sphincter of Oddi dysfunction; STRI, strictures; n/a, not available.
isocitrate and hippurate and a higher level of sulfated glycol-ursodeoxycholic acid was characteristic of the urinary metabotype of patients with bile duct strictures, compared to healthy controls. Only one molecule, isocitrate, achieved statistical significance between patients with strictures and healthy controls.

The urinary metabolites altered between benign and malignant bile duct strictures were primarily related to

---

**Figure 1** CCA patients vs healthy controls. PCA scores plots for (A) positive and (B) negative ion mode data of CCA patients and healthy controls. OPLS-DA scores plots showing group separation for both (C) positive and (D) negative ion mode data and the corresponding permutation tests for (E) positive and (F) negative ion mode data.

**Abbreviations:** CCA, cholangiocarcinoma; PCA, principal component analysis; OPLS-DA, orthogonal projections to latent structures discriminant analysis.
acylcarnitine metabolism, where the levels of (C9-OH, C10 and C10:1) carnitines were significantly downregulated in subjects with benign strictures. 

Definitive diagnosis to distinguish benign from malignant lesions involves multiple tools including imaging, endoscopic investigations, tissue sampling
and diagnostic surgery.\(^7\) Even after preoperative evaluation, biliary strictures can remain indeterminate.\(^8\) The serum metabolic signatures of PSC patients in US population have been profiled using LC and GC mass-spectroscopic detection techniques.\(^9\) Markers related to inflammatory status, as well as to
Perturbations in bile acid metabolism (elevations in cholesterol, unconjugated, conjugated and sulfated bile acid), lipid metabolism (elevations in fatty acids, ketone bodies and several acylcarnitines) and protein metabolism (depletion in dipeptides), were associated with the PSC patient’s metabolome, compared to healthy controls. The authors postulated that these changes reflect cholestatic liver state, perturbations in

**Figure 4** CCA patients vs biliary strictures patients. PCA scores plots for (A) positive and (B) negative ion mode data of CCA patients and biliary strictures patients. OPLS-DA scores plots showing group separation for both (C) positive and (D) negative ion mode data and the corresponding permutation tests for (E) positive and (F) negative ion mode data.

**Abbreviations:** CCA, cholangiocarcinoma; PCA, principal component analysis; OPLS-DA, orthogonal projections to latent structures discriminant analysis.
cholesterol homeostasis and mitochondrial dysfunction which are typical features associated with impaired fatty acid $\beta$-oxidation.

It is not clear why the urine metabolome of patients with benign biliary strictures was distinguishable from healthy controls, but not from those with malignant biliary

---

**Figure 5** CCA vs noncirrhotic HCV patients. PCA scores plots for (A) positive and (B) negative ion mode data of CCA patients and HCV patients. OPLS-DA scores plots showing group separation for both (C) positive and (D) negative ion mode data and the corresponding permutation tests for (E) positive and (F) negative ion mode data.

**Abbreviations:** CCA, cholangiocarcinoma; HCV, noncirrhotic hepatitis C; PCA, principal component analysis; OPLS-DA, orthogonal projections to latent structures discriminant analysis.
tumors. These results may suggest that cellular turnover related to underlying biliary tract inflammation in benign or malignant cholestatic liver disease is likely to exhibit similar metabolic traits in the urine metabolome. However, it is difficult to draw firm conclusions from these data due to the small sample size (CCA=10 participants and benign biliary strictures=9 participants).

In the current literature, metabolic phenotyping studies in hepatobiliary disease have been mainly focused on bile profiling. Differences in the biochemical composition of bile between CCA and patients with nonmalignant biliary disease (including various conditions such as SOD, PSC and gallstones) were distinguished between the two groups.\textsuperscript{10,11} One study examined bile from PSC (n=18) patients and from patients with CCA (n=16), but also included 11 who had other benign biliary conditions. The spectra of bile from CCA patients differed from the benign group in the levels of phosphatidylcholine, bile acids, lipids and cholesterol with sensitivity, specificity and accuracy of 88.9%, 87.1% and 87.8%, respectively.\textsuperscript{12}

### Table 2: Altered metabolites between cholangiocarcinoma patients compared to healthy controls – positive mode

<table>
<thead>
<tr>
<th>m/z</th>
<th>RT</th>
<th>Tentative assignment</th>
<th>Adduct</th>
<th>VIP</th>
<th>Trend</th>
<th>p-value\textsuperscript{†}</th>
<th>FC</th>
<th>Identification*</th>
</tr>
</thead>
<tbody>
<tr>
<td>162.113</td>
<td>0.49</td>
<td>L-carnitine (C1)</td>
<td>M+H</td>
<td>4.3</td>
<td>↑</td>
<td>NS</td>
<td>2.62</td>
<td>a</td>
</tr>
<tr>
<td>114.066</td>
<td>0.51</td>
<td>Creatinine</td>
<td>M+H</td>
<td>5.9</td>
<td>↓</td>
<td>0.043</td>
<td>−1.4</td>
<td>a</td>
</tr>
<tr>
<td>144.102</td>
<td>0.58</td>
<td>Proline betaine</td>
<td>M+H</td>
<td>4.5</td>
<td>↓</td>
<td>NS</td>
<td>−0.67</td>
<td>b</td>
</tr>
<tr>
<td>229.118</td>
<td>0.68</td>
<td>Hydroxyproplyl-proline</td>
<td>M+H</td>
<td>3.3</td>
<td>↑</td>
<td>NS</td>
<td>1.94</td>
<td>b</td>
</tr>
<tr>
<td>204.125</td>
<td>0.82</td>
<td>L-Acetylcarnitine (C2)</td>
<td>M+H</td>
<td>7.7</td>
<td>↑</td>
<td>NS</td>
<td>1.87</td>
<td>a</td>
</tr>
<tr>
<td>229.155</td>
<td>0.98</td>
<td>Leucyl-proline</td>
<td>M+H</td>
<td>5.2</td>
<td>↑</td>
<td>NS</td>
<td>1.45</td>
<td>b</td>
</tr>
<tr>
<td>169.036</td>
<td>1.03</td>
<td>Uric acid</td>
<td>M+H</td>
<td>3</td>
<td>↓</td>
<td>NS</td>
<td>−0.9</td>
<td>a</td>
</tr>
<tr>
<td>215.016</td>
<td>1.06</td>
<td>Citrate</td>
<td>M+Na</td>
<td>3.2</td>
<td>↓</td>
<td>NS</td>
<td>−3.58</td>
<td>a</td>
</tr>
<tr>
<td>188.071</td>
<td>2.51</td>
<td>Unknown</td>
<td>-</td>
<td>4.8</td>
<td>↑</td>
<td>NS</td>
<td>2.22</td>
<td>d</td>
</tr>
<tr>
<td>232.155</td>
<td>2.84</td>
<td>Butyrylcarnitine (C4)</td>
<td>M+H</td>
<td>4.4</td>
<td>↓</td>
<td>0.023</td>
<td>−2.22</td>
<td>a</td>
</tr>
<tr>
<td>296.15</td>
<td>3.17</td>
<td>Unidentified acylcarnitine</td>
<td>M+H</td>
<td>3.3</td>
<td>↑</td>
<td>NS</td>
<td>0.14</td>
<td>a</td>
</tr>
<tr>
<td>295.13</td>
<td>3.54</td>
<td>Glutamylphenylalanine</td>
<td>M+H</td>
<td>3.2</td>
<td>↑</td>
<td>&lt;0.0001</td>
<td>12.19</td>
<td>b</td>
</tr>
<tr>
<td>246.17</td>
<td>3.57</td>
<td>Valerylcaritine (C5)</td>
<td>M+H</td>
<td>3.4</td>
<td>↓</td>
<td>NS</td>
<td>−1.78</td>
<td>a</td>
</tr>
<tr>
<td>318.191</td>
<td>3.78</td>
<td>Acylcarnitine (C9-OH)</td>
<td>M+H</td>
<td>6.3</td>
<td>↑</td>
<td>&lt;0.0001</td>
<td>9.25</td>
<td>b</td>
</tr>
<tr>
<td>178.05</td>
<td>3.8</td>
<td>Hippurate</td>
<td>M+H</td>
<td>8.9</td>
<td>↓</td>
<td>&lt;0.0001</td>
<td>−3</td>
<td>a</td>
</tr>
<tr>
<td>344.207</td>
<td>4.4</td>
<td>Acylcarnitine (C10:1)</td>
<td>M+H</td>
<td>3.1</td>
<td>↑</td>
<td>0.001</td>
<td>4.43</td>
<td>b</td>
</tr>
<tr>
<td>346.223</td>
<td>4.57</td>
<td>Acylcarnitine (C10)</td>
<td>M+H</td>
<td>3.9</td>
<td>↑</td>
<td>0.0002</td>
<td>7.88</td>
<td>b</td>
</tr>
<tr>
<td>170.061</td>
<td>4.81</td>
<td>Unknown</td>
<td>-</td>
<td>5.2</td>
<td>↓</td>
<td>0.001</td>
<td>−3.41</td>
<td>d</td>
</tr>
<tr>
<td>286.201</td>
<td>4.92</td>
<td>Acylcarnitine (C8:1)</td>
<td>M+H</td>
<td>6.4</td>
<td>↓</td>
<td>NS</td>
<td>−0.79</td>
<td>b</td>
</tr>
<tr>
<td>277.142</td>
<td>5.12</td>
<td>Unknown</td>
<td>M+H</td>
<td>3.5</td>
<td>↑</td>
<td>NS</td>
<td>0.1</td>
<td>b</td>
</tr>
<tr>
<td>312.217</td>
<td>5.16</td>
<td>2-trans,4-cis-Decadienoylearnitine (C10)</td>
<td>M+H</td>
<td>3.9</td>
<td>↓</td>
<td>NS</td>
<td>−5.78</td>
<td>b</td>
</tr>
<tr>
<td>310.202</td>
<td>5.19</td>
<td>Acylcarnitine (C10:3)</td>
<td>M+H</td>
<td>5.2</td>
<td>↓</td>
<td>NS</td>
<td>−1.57</td>
<td>b</td>
</tr>
<tr>
<td>310.202</td>
<td>5.35</td>
<td>Acylcarnitine (C10:3)</td>
<td>M+H</td>
<td>5.4</td>
<td>↑</td>
<td>0.021</td>
<td>−1.85</td>
<td>b</td>
</tr>
<tr>
<td>302.233</td>
<td>5.55</td>
<td>2,6-Dimethylheptanoyl carnitine (C9)</td>
<td>M+H</td>
<td>7.6</td>
<td>↓</td>
<td>0.001</td>
<td>3.65</td>
<td>b</td>
</tr>
<tr>
<td>314.233</td>
<td>5.78</td>
<td>Decenoylarnitine (C10:1)</td>
<td>M+H</td>
<td>3.1</td>
<td>↑</td>
<td>NS</td>
<td>1.62</td>
<td>b</td>
</tr>
<tr>
<td>328.248</td>
<td>6.11</td>
<td>Acylcarnitine (C102-OH)</td>
<td>M+H</td>
<td>4</td>
<td>↓</td>
<td>NS</td>
<td>1.43</td>
<td>b</td>
</tr>
</tbody>
</table>

Notes: *Level of metabolite identification: (a) identified compound; (b) putatively annotated compound; (c) putatively characterized compound class; and (d) unknown. †False discovery rate (FDR) adjusted p-value (or q-value).

Abbreviations: FC, fold change; RT, retention time; VIP, variable importance in projection score.

**Metabolic profile discrimination between CCA and malignant disease groups**

The metabolic profile acquired using urine specimens from CCA patients was distinguishable from those with hepatic and ovarian malignancies, but not from individuals with metastatic and pancreatic tumors. Urinary acylcarnitine pattern and excess excretion of urinary bile acids were comparable between the two groups. The urine profiles from pancreatic cancer patients were not distinguishable from CCA and metastatic patients and also shared similarity, based on their urinary biochemical composition.

Pancreatic and bile duct tumors (particularly bile duct tumors that arise in the distal extrahepatic region) share a close anatomical relation, similar growth pattern and close phenotypic characteristics and are associated with poor prognosis, which may explain the similarity in their
In a recently published review on pancreatic cancer, unlike CCA, it is not rare: it is a rising cause of cancer death in Europe. The predicted death trend in both genders, 4% in men and 5% in women between 2009 and 2015, corresponding to 85,300 total deaths in 2015. It is difficult to draw a complete picture of the biological pathways implicated in pancreatobiliary conditions from this small dataset. To enable accurate metabolomic characterization of pancreatobiliary tumors, metabolic fingerprinting using large sample size and more complete clinical information is needed. The possible application of metabolic profiling to differentiate different tumor localization and disease staging would helpfully provide a more comprehensive picture of tumor development and progression in the tumors of the pancreatobiliary junction.

Table 3 Altered metabolites between cholangiocarcinoma patients compared to healthy controls – negative mode

| m/z | RT | Tentative assignment | Adduct | VIP | Trend | p-value | FC | Identification
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>96.962</td>
<td>0.60</td>
<td>Sulfate</td>
<td>M-H</td>
<td>3.8</td>
<td>↓</td>
<td>0.025</td>
<td>-4.83</td>
<td>b</td>
</tr>
<tr>
<td>243.061</td>
<td>0.85</td>
<td>Uridine/pseudouridine</td>
<td>M-H</td>
<td>3.1</td>
<td>↑</td>
<td>0.0002</td>
<td>2.24</td>
<td>b</td>
</tr>
<tr>
<td>191.018</td>
<td>0.94</td>
<td>Isocitrate</td>
<td>M-H</td>
<td>5.8</td>
<td>↓</td>
<td>0.023</td>
<td>-1.58</td>
<td>b</td>
</tr>
<tr>
<td>167.02</td>
<td>1.03</td>
<td>Uric acid</td>
<td>M-H</td>
<td>3.1</td>
<td>↑</td>
<td>0.011</td>
<td>1.7</td>
<td>a</td>
</tr>
<tr>
<td>191.018</td>
<td>1.06</td>
<td>Citrate</td>
<td>M-H</td>
<td>4.7</td>
<td>↓</td>
<td>NS</td>
<td>-1.93</td>
<td>a</td>
</tr>
<tr>
<td>227.996</td>
<td>2.65</td>
<td>Sulfated compound</td>
<td>M-H</td>
<td>3.2</td>
<td>↓</td>
<td>0.005</td>
<td>-6.12</td>
<td>d</td>
</tr>
<tr>
<td>188.985</td>
<td>3.13</td>
<td>Pyrocatechol sulfate</td>
<td>M-H</td>
<td>4.5</td>
<td>↓</td>
<td>0.03</td>
<td>-2.66</td>
<td>b</td>
</tr>
<tr>
<td>172.991</td>
<td>3.45</td>
<td>Phenyl sulfate</td>
<td>M-H</td>
<td>6.8</td>
<td>↓</td>
<td>0.039</td>
<td>-3.44</td>
<td>b</td>
</tr>
<tr>
<td>203.001</td>
<td>3.75</td>
<td>O-methoxycatechol-O-sulfate</td>
<td>M-H</td>
<td>3.0</td>
<td>↓</td>
<td>0.014</td>
<td>-2.45</td>
<td>b</td>
</tr>
<tr>
<td>178.049</td>
<td>3.80</td>
<td>Hippurate</td>
<td>M-H</td>
<td>8.5</td>
<td>↓</td>
<td>0.006</td>
<td>-2.2</td>
<td>a</td>
</tr>
<tr>
<td>263.102</td>
<td>3.84</td>
<td>Phenylacetylglutamine</td>
<td>M-H</td>
<td>4.4</td>
<td>↓</td>
<td>NS</td>
<td>1.11</td>
<td>b</td>
</tr>
<tr>
<td>212.001</td>
<td>3.87</td>
<td>Indoxylysufate</td>
<td>M-H</td>
<td>9.1</td>
<td>↓</td>
<td>0.021</td>
<td>-1.93</td>
<td>a</td>
</tr>
<tr>
<td>241.118</td>
<td>3.92</td>
<td>L-gamma-glutamyl-L-isoleucine</td>
<td>M-H</td>
<td>3.2</td>
<td>↑</td>
<td>0.008</td>
<td>3.14</td>
<td>b</td>
</tr>
<tr>
<td>287.022</td>
<td>4.13</td>
<td>DHPV O-sulfate</td>
<td>M-H</td>
<td>4.2</td>
<td>↓</td>
<td>0.018</td>
<td>-6.32</td>
<td>b</td>
</tr>
<tr>
<td>243.134</td>
<td>4.18</td>
<td>Hydroxypoeryl-isoleucine</td>
<td>M-H</td>
<td>2.9</td>
<td>↓</td>
<td>0.002</td>
<td>3.24</td>
<td>b</td>
</tr>
<tr>
<td>187.005</td>
<td>4.41</td>
<td>p-cresol sulfate</td>
<td>M-H</td>
<td>1.7</td>
<td>↓</td>
<td>0.051</td>
<td>-1.85</td>
<td>b</td>
</tr>
<tr>
<td>352.085</td>
<td>5.88</td>
<td>Fatty acid</td>
<td>M-H</td>
<td>3.6</td>
<td>↓</td>
<td>0.004</td>
<td>-5.31</td>
<td>d</td>
</tr>
<tr>
<td>465.248</td>
<td>6.37</td>
<td>Steroid glucuronide (C25H38O8)</td>
<td>M-H</td>
<td>3.0</td>
<td>↓</td>
<td>0.001</td>
<td>3.9</td>
<td>b</td>
</tr>
<tr>
<td>419.228</td>
<td>6.48</td>
<td>Steroid compound</td>
<td>M-H</td>
<td>3.1</td>
<td>↓</td>
<td>0.005</td>
<td>15.3</td>
<td>d</td>
</tr>
<tr>
<td>528.263</td>
<td>6.76</td>
<td>Glycocoluxinlysosylchol-N-sulfate</td>
<td>M-H</td>
<td>4.4</td>
<td>↑</td>
<td>0.011</td>
<td>18.3</td>
<td>b</td>
</tr>
</tbody>
</table>

Notes: *Level of metabolite identification: (a) identified compound; (b) putatively annotated compound; (c) putatively characterized compound class; and (d) unknown. False discovery rate (FDR) adjusted p-value (or q-value). 5′-(3′,4′-dihydroxyphenyl)-gamma-valerolactone sulfate.

Abbreviations: FC, fold change; RT, retention time; VIP, variable importance in projection score.

Metabolic dysfunctions in several molecular pathways implicated in carcinogenesis such as ketogenesis (increase in acetoacetate and decrease in citrate) were described to be associated with the nuclear magnetic resonance urine metabolome from pancreatic cancer patients. Napoli et al investigated the spectral correlations, in the urine metabolic profile from individuals with pancreatic cancer, with pathological staging and tumor anatomical localization. Patients with intermediate pancreatic carcinoma stage showed different urinary metabolome from those with advanced tumor stage. Also, a clear distinction was observed based on the anatomical location of pancreatic cancer, particularly between carcinoma of the uncinate process of the pancreas and pancreatic cancers of the body and head. Unlike CCA, pancreatic cancer is not rare: it is the fourth cause of cancer death in Europe. It has a rising predicted death trend in both genders, 4% in men and 5% in women between 2009 and 2015, corresponding to 85,300 total deaths in 2015. It is difficult to draw a complete picture of the biological pathways implicated in pancreatobiliary conditions from this small dataset. To enable accurate metabolomic characterization of pancreatobiliary tumors, metabolic fingerprinting using large sample size and more complete clinical information is needed. The possible application of metabolic profiling to differentiate different tumor localization and disease staging would helpfully provide a more comprehensive picture of tumor development and progression in the tumors of the pancreatobiliary junction.
Compared to HCC and OCA cases, urine from CCA patients showed a distinct pattern. A number of metabolites were significantly different between the urine profile from CCA patients, compared to HCC and OCA cases and may possibly serve as candidate biomarkers in CCA. These include acylcarnitine (C9:OH), L-carnitine (C1), valerylcarnitine (C5), acylcarnitine (C8:1), 2,6 dimethylheptanoylcarnitine (C9), L-acetylcarnitine (C2), 12-trans, 4-cis-Decadienoylcarnitine (C10), 2-trans, 4-cis-Decadienoylcarnitine (C10), 2-trans, 4-cis-Decadienoylcarnitine (C10), and 2-trans, 4-cis-Decadienoylcarnitine (C10).

Figure 6 Box and whisker plots of compounds different between cholangiocarcinoma (CCA) cases and benign controls – carnitine species. *p 0.05, **p 0.01 and ***p 0.001.

Y-axis: relative intensity.

Abbreviations: CCA, cholangiocarcinoma; GS, gallstones; SOD, sphincter of Oddi dysfunction; STRI, strictures; HCV, noncirrhotic hepatitis C.
Uridine/pseudouridine, glucosamine-6-phosphate and certain bile acids. The key metabolic pathways most associated with cholangiocarcinogenesis are discussed below.

**Acylcarnitine metabolism**

Several species of acylcarnitines were dysregulated in individuals with CCA, compared to control groups. An acylated carnitine compound, tentatively identified as acylcarnitine C9-OH, was significantly greater abundance in CCA metabolic profiles compared to all groups, except patients with pancreatic and metastatic tumors. Multigroup PLS-DA identified carnitine-related metabolites (carnitine [C1], acetylcarnitine [C2] and acylcarnitine [C9-OH]) to be most associated with CCA cases, compared to nonmalignant conditions.

Elevated urinary concentrations of acylcarnitine metabolites were also observed in individuals with gallstones. ANOVA analysis showed that individuals with gallstones had significantly greater levels of acylcarnitine metabolites (including C2, C9-OH, C10 and C10:1), with respect to healthy controls. Gallstone disease is a common disorder of the hepatobiliary system, characterized by the formation of gallstones (or cholelithiasis) in the gallbladder, common bile duct or hepatic bile duct. Cholelithiasis is a well-established risk factor for...
gallbladder carcinoma; it causes severe injury to the biliary mucosa which triggers a series of events starting with chronic inflammation to metaplasia, dysplasia, in situ carcinoma to invasive carcinoma.19

Gallbladder stone chemical composition is heterogeneous and the stone formation is aggravated by several lifestyle and environmental factors, such as low physical activity, high dietary fat intake, female gender

Figure 8 Box and whisker plots of compounds different between cholangiocarcinoma (CCA) cases and benign controls: ESI− metabolites. *p 0.05, **p 0.01 and ***p 0.001. Y-axis: relative intensity.

Abbreviations: CCA, cholangiocarcinoma; GS, gallstones; SOD, sphincter of Oddi dysfunction; STRI, strictures; HCV, noncirrhotic hepatitis C.
and genetic predisposition.\textsuperscript{20} The presence of metabolic disorders including adiposity, hypertension, diabetes mellitus and lipid abnormalities is also associated with a higher prevalence of gallstone disease.\textsuperscript{20} Considering all these factors, in addition to the lack of comprehensive metabolic profiling studies on the disease, makes it difficult to attempt to capture and characterize the urinary biochemical perturbation in individuals with cholelithiasis. Impaired metabolism of cholesterol, phospholipids, bilirubin and bile acids is known to be implicated in the pathogenesis of gallstone disease.\textsuperscript{19}

Metabolism of pyrimidine and purine nucleotides

Nucleotides, purines and pyrimidines are the nitrogenous bases of the genetic code.\textsuperscript{21} Nucleotides and their derivatives contribute to many biological processes such as cellular signaling, energy production and lipid and protein synthesis. Aberrant DNA replication and damage occur early during cellular carcinogenesis, in which imbalance of nucleotide metabolism, particularly deoxyribonucleotide triphosphates (dNTP), drives tumorigenesis and leads to nucleotide pool imbalance in tissue and biological...
fluids. These nucleotide pool patterns may serve as a novel screening and diagnostic biomarkers for various human cancers. Increased concentrations of modified nucleotides (such as pseudouridine, 1-methyladenosine and 1-methylguanosine) in the urine of cancer patients are reflective of high whole-body RNA turnover or oxidative DNA damage.

Nucleotide post-transcriptional chemical modifications, including methylation, hydroxylation, acetylation and uridine isomerization, are believed to play a key role in the translation of the genetic code, yet their exact biological function remains unclear. Ridine or pseudouridine, a structural C-glycoside isomer of the nucleoside uridine, was significantly increased in CCA compared to all groups except individuals with strictures and pancreatic tumors. The pyrimidine nucleoside plays a crucial role in the synthesis of RNA, glycogen and biomembrane. Uridine is used clinically as a rescue agent that protects against 5-fluorouracil toxicity. Medication history records showed that none of the individuals with CCA consumed 5-fluorouracil or uridine-containing prodrugs.

The circulating plasma uridine level is tightly regulated in humans, yet several factors can result in elevated plasma uridine concentration, such as enhanced ATP consumption, enhanced uridine diphosphate (UDP)-glucose consumption via glycogenesis, increased urea synthesis and increased 5-phosphoribosyl-1-pyrophosphate. Ka et al demonstrated positive correlations between plasma uridine concentrations and the urinary excretions of urea, uric acid, uridine, uric acid clearance and purine intake in healthy males. The liver regulates and maintains the hepatic pools of uridine nucleotides via de novo uridine synthesis and degradation which is essential for the homeostatic control of plasma uridine pools. An in vivo study evaluated uridine homeostasis in liver tissue and its impact on hepatic cellular function. The authors revealed that uridine administration suppresses fatty liver by modulating liver protein acetylation profile and identified an association between uridine homeostasis, pyrimidine metabolism and liver lipid metabolism.

We found an imbalance of purine metabolites in the urinary metabolome of CCA patients. Purine-based metabolites in this UK cohort showed increased uric acid, increased 7-methylguanine and decreased hypoxanthine. In the discriminant analysis applied to generate metabolite patterns as a method of disease identification, uric acid, hypoxanthine and 7-methylguanine were all key metabolites responsible for the differences in metabolic signatures in CCA cases compared to the malignant controls.

**Bile acid metabolism**
The greatest abundance of bile acid species was found in the urine metabolic profiles from individuals with pancreatobiliary tumors. It is difficult to assess bile acid profile in the bile duct and pancreatic carcinomas as they are frequently complicated with profound jaundice, which subsequently causes...
Cholangiocarcinoma patients vs. hepatocellular carcinoma

A marked increase in bile acid concentrations in both serum and urine. Nevertheless, elevation in bile acid species, such as tauroursodeoxycholic acid, taurocholic acid, deoxycholylglycine and cholylglycine, has been observed in the plasma MS metabolic profiles from nonjaundiced patients with early-stage pancreatic ductal adenocarcinoma. In a recent metabonomics study by Di Gangi et al, free plasma bile acid concentrations were significantly decreased, whereas conjugated (glycine- and taurine-) cholic and chenodeoxycholic acids were significantly increased in

Figure 11 CCA patients vs HCC patients. PCA scores plots for (A) positive and (B) negative ion mode data of CCA patients and HCC patients. OPLS-DA scores plots showing group separation for both (C) positive and (D) negative ion mode data and the corresponding permutation tests for (E) positive and (F) negative ion mode data. Abbreviations: CCA, cholangiocarcinoma; HCC, hepatocellular carcinoma; PCA, principal component analysis; OPLS-DA, orthogonal projections to latent structures discriminant analysis.
pancreatic cancer cases compared to disease-free controls. Cytotoxicity and cytoprotection by bile acids vary depending on their hydrophobic/hydrophilic properties. In jaundiced serum of pancreatic cancer cases, elevation in conjugated bile acids seems to exhibit a protective effect against pancreatic cell proliferation. Possible mechanisms of the inhibitory action of bile acids involve their cytotoxic potency which reduces pancreatic cancer cell line proliferation and results in structural damage. Little is known about the causative role in metabolic perturbations and oncogenic pathways preceding pancreatic cancer progression.

In their review, Feng et al aimed to evaluate bile acids role in the etiology of pancreatic carcinogenesis. However, more questions are raised than have been answered; bile acids are not only involved in multiple risk factors known to be implicated in pancreatic cancer initiation, including

**Figure 12** CCA patients vs ovarian cancer. PCA scores plots for (A) positive and (B) negative ion mode data of CCA patients and OCA patients. OPLS-DA scores plots showing group separation for both (C) positive and (D) negative ion mode data and the corresponding permutation tests for (E) positive and (F) negative ion mode data. **Abbreviations:** PCA, principal component analysis; CCA, cholangiocarcinoma; OCA, ovarian cancer; OPLS-DA, orthogonal projections to latent structures discriminant analysis.
Obesity, diabetes, high-fat diet and gallstones, but they also exhibit biologically complex local tissue effects. They play a pathogenic and/or protective role which is complex to simulate the ex vivo environment.

Conclusions
Primary data from this dataset identified multimolecular signatures of pancreatobiliary disease and illustrated the potential of MS-based metabolic profiling in generating a novel noninvasive diagnostic tool for tumor detection and may further understand disease mechanisms. Discriminant metabolites related to biliary stricturing conditions (both benign and CCA) showed a similar pattern. It was not possible to differentiate between the metabolic profiles of patients with pancreatic cancer from CCA or from premalignant biliary strictures, but metabolic profiles of patients with HCC and OCAs were distinct from those with pancreatobiliary pathology.

It is important to highlight that its difficult to draw firm conclusions from the limited sample size. In-depth metabolic characterization of biofluids from patients with a range of pancreatobiliary conditions is required to further characterize metabonomic signatures for CCA.

Acknowledgments
This study was funded by grants from the Welcome Trust ISSF Fund at Imperial College London and AMMF – the Cholangiocarcinoma Charity (Stansted, Essex, UK). MA was funded by the StratiGrad PhD programme at Imperial College London. All authors acknowledge the support of the United Kingdom National Institute for Health Research Biomedical Research Centre at Imperial College London for infrastructure support. We would like to thank Dr Shaun Greer and Dr Martin Prince of Manchester Royal Infirmary Centre for Hepatology, as well as Dr I. Jane Cox of the Institute of Hepatology at King’s College London for professional collaboration on sample collection and on study design. We would also like to show our gratitude to Professors Paiboon Sithithaworn, Narong Khuntikeo, Ross H. Andrews, Watcharin Loilome and Puangrat Yongvanit of Khon Kaen University in Thailand for their insight and expertise. Finally, we would also like to thank Larry K. Koomson, Abigail Zabron, Mary M. E. Crossey and Christopher A. Wadsworth of Imperial College London for their technical and administrative support.

Disclosure
The authors report no conflicts of interest in this work.

References


