SUPPLEMENTARY MATERIAL:

MATERIAL AND METHODS:

Study Participants and Procedures:

The exclusion criteria included: those with hepatitis B or C, autoimmune hepatitis, primary biliary cirrhosis, primary sclerosing cholangitis, Wilson disease, hemochromatosis, α-1 antitrypsin deficiency, average alcohol intake greater than 20 gm/day (women) and >30 gm/day (men) , medications known to cause hepatic steatosis such as prednisone, amiodarone etc. or history of bowel surgery. Also, subjects taking pharmacological doses of ursodeoxycholic acid, cholestyramine or colsevelam for any reason were excluded.

All subjects were seen at the Virginia Commonwealth University Medical Center's General Clinical Research Unit. Routine history and physical exam was performed and pertinent data were recorded. Fasting blood samples were drawn for biochemical tests and additional samples were collected into an EDTA-containing tube for bile acid analysis. Blood samples were centrifuged at 4°C, and the plasma was stored at −80°C. Liver biopsy was performed on subjects with suspected NAFLD.

The liver histology was assessed by two dedicated hepato-pathologists; of these, one helped develop the NASH Clinical Research Network (CRN) scoring system and is a senior experienced pathologist. The presence of fatty liver disease (> 5% steatosis) and its phenotype (fatty liver, borderline or definite steatohepatitis) were recorded along with the severity of individual histological features of NASH such as steatosis, inflammation, hepatocellular ballooning and fibrosis measured by the NASH CRN pathology committee protocol (1). The activity of the disease was quantified by the

NAFLD activity score (NAS) while the stage of the disease was imputed from the fibrosis stage.

In controls, age, aspartate aminotransferase (AST), alanine aminotransferase (ALT) and platelet count were used to calculate a FIB-4 score for NASH as previously described (2) for staging of fibrosis: F0-F1 (≤1.30) and F3-F4 (≥2.67). NAFL was defined as NAFLD with a NAFLD Activity Score (NAS) ≤3 along with an absence of hepatocellular ballooning. Subjects with active definite NASH were included in the study. This was defined by the presence of definite steatohepatitis and a NAS ≥4.

Bile Acid profiling studies:

Sample Preparation: The plasma sample was prepared using the automated MicroLab STAR® system as previously described (3). Recovery standards were added prior to the first step in the extraction process and a series of organic and aqueous extractions were performed to remove the protein fraction while allowing maximum recovery of small molecules for quality assurance and quality control. The resulting extract was used for analysis by LC. Samples were placed briefly on a TurboVap® (Zymark) to remove the organic solvent. Each sample was then frozen and dried under vacuum. Samples were then prepared for the LC/MS.

Liquid chromatography/Mass Spectrometry (LC/MS): The LC/MS portion of the platform was based on a Waters ACQUITY UPLC and a Thermo-Finnigan LTQ mass spectrometer, which consisted of an electrospray ionization (ESI) source and linear ion-trap (LIT) mass analyzer. The sample extract was split into two aliquots, dried, then reconstituted in acidic or basic LC-compatible solvents, each of which contained 11 or

more injection standards at fixed concentrations. One aliquot was analyzed using acidic positive ion optimized conditions and the other using basic negative ion optimized conditions in two independent injections using separate dedicated columns. Extracts reconstituted in acidic conditions were gradient eluted using water and methanol both containing 0.1% Formic acid, while the basic extracts, which also used water/methanol, contained 6.5mM Ammonium Bicarbonate. Following LC/MS runs, the metabolites were identified based on the combination of chromatographic and mass spectra properties by automated comparison to the metabolomic library entries of purified standards. Batch normalization was performed using the median ratio for each metabolite in duplicate "anchor" samples across runs.

Enzyme-linked immunosorbent assay (ELISA)

An ELISA kit was used for colorimetric detection of human FGF19 present in plasma of controls (n=8), NAFL (n=8), and NASH (n=24, 8 F0-F1, 8 F2 and 8 F3-F4) subjects according to the manufacturer's instructions (Human FGF-19 ELISA Kit (EHFGF19), Thermo Scientific). All plasma samples were assayed in duplicate and one way analysis of variance (ANOVA) was performed with the significance level set at P-value <0.05.

RNA Isolation and Quantitative RT-PCR.

Human liver tissues from control (n=6), NAFL (n=6) and NASH (n=6) subjects were lysed in RLT lysis buffer as followed by manufacturer's instructions (QIAGEN, 74104). Total RNA yield was measured by using a Nano Drop spectrophotometer (DeNovix). Total RNA isolated was reverse transcribed into cDNA by using random primers (Invitrogen) and reverse transcriptase enzyme (MLV, Invitrogen). Quantification of the cDNA template

was performed by real-time PCR using SYBR green fluorescence on a Light Cycler 480 instrument (Roche Applied Science). All gene specific primers were purchased from Thermo Fisher Scientific and with the following primer sequences: human SHP (NR0B2): forward 5'-CTTCTGGAGCCTGGAGCTTA-3' and reverse 5'-ACTTCACACAGCACCCAGTG-3'; human GAPDH: forward 5'-CTTCACCACCATGGAGGAGGC-3' and reverse 5'-GGCATGGACTGTGGTCATGAG-3'; human Cyp7A1: forward 5'-CCATTAGGTGTTGTGCCACG-3' and reverse 5'-CATCCATCGGGTCAATGCTT-3'; human PLPT: forward 5'-GCTCTACTGGTTCTTCTATG -3' and reverse 5'- TCAATGCCAACAAGCTCGTC -3' and human BSEP (ABCB11): forward 5'- TGAGCCTGGTCATCTTGTG-3' and reverse 5'-TCCGTAAATATTGGCTTTCTG -3'. The relative quantity of mRNA was expressed relative to the amount of the reference GAPDH gene according to the comparative threshold cycle (Ct) method.

Bioinformatics and statistical analysis:

Descriptive statistics and frequency distributions were generated for the sample demographic and clinical characteristics. Group means and medians were calculated and statistically significant differences ($p<0.05$) identified using analysis of variance (ANOVA) followed by Tukey's post-hoc pairwise tests. Nonparametric tests were performed for non-normally distributed with Dunn’s post hoc reporting multiplicity adjusted P-value for each comparison. Logistic regression analyses were performed for the association between BA metabolites and disease phenotypes as well as histologic features. Logistic regression analyses model included all major bile acids as predictors with various histologic features as outcomes: nominal logistics for dichotomous (steatosis ≤66% vs. >66%, NAFLD

Activity Score (NAS) ≤3 vs. ≥4, and fibrosis F0-F1 vs. ≥2), ordinal logistics for multinomial (lobular and portal inflammation) and least square regression analysis for continuous (NAS from 0-8). The predictors were removed in step wise manner until all but the most statistically significant predictors remained. For multiple comparisons in the regression models false discovery rate (FDR) LogWorth for each model effect, defined as -log10(FDR PValue) was calculated for the association between plasma BAs and histologic features as appropriate. The mean of log-transformed BA measurements for each respective group was displayed as heat maps. The colors in the heat map display indicate the mean of the bile acid measurements for each respective group. The measurements are on a log-scale with smaller measurements indicated by blues/greens and larger measurements indicated by oranges/reds. Yellow values are in between. Negative log values indicate untransformed values that are less than 1 and positive log values indicate untransformed values that are greater than 1. The side bar along the left side of the heat map is color coded for black, yellow, or green. Black indicates those BAs with differences that are not statistically significant across groups. Yellow indicates those BAs with differences with p-values <0.1 but greater than the statistical significance threshold of $p < 0.05$. Green indicates those BAs that do meet the statistical significance threshold of $p < 0.05$. These statistical analyses were performed using JMP, Graph Pad, and “R” base(4), multcomp (5), and PMCMR packages (6) http://cran.r-project.org/.

REFERENCES:

1. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, Ferrell LD, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313-1321.

2. Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ, Nash Clinical Research N. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2009;7:1104-1112.

3. Evans AM, DeHaven CD, Barrett T, Mitchell M, Milgram E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. Anal Chem 2009;81:6656-6667.

4. R: A language and environment for statistical computing. 2014.

5. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. Biom J 2008;50:346-363.

6. Pohlert T. The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR). In; 2014.

Supplementary Table 1: **(A)** Clinical and demographic profile of study population. **(B)** Clinical and demographic characteristics of study population according to different stages of fibrosis. ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; F0/F1, stage 0/stage1 fibrosis; F≥2, stage 2 or greater fibrosis.

Supplementary Figure 1: **(A)** No statistically significant difference was observed in the plasma 7-alpha-hydroxyl-3-oxo-4-cholestenoic acid (C4) levels among the study groups. Similarly, no statistically significant differences among the study groups in **(B)** cholate, **(C)** chenodeoxycholate (CDCA), **(D)** primary unconjugated BAs, **(E)** conjugated CDCA.

Supplementary Figure 2: No statistically significant differences were noted among the study groups in **(A)** glycodeoxycholate (GDCA), **(B)** taurodeoxycholate (TDCA), **(C)** glycolithocholate (GLCA), **(D)** taurolithocholate (TLCA), **(E)** total conjugated deoxycholate (DCA), and **(F)** total conjugated secondary BAs.

Supplementary Figure 3: **(A)** Higher NAFLD Activity Score (NAS) ≥4 is significantly associated with increased taurocholate. **(B)** Glycocholate is significantly higher in subjects with higher stages (F≥2) of fibrosis. **(C)** F≥2 fibrosis is associated with statistically significant higher taurocholate. *$p<0.05$, **$p<0.01$.

Supplementary Figure 4: **(A)** The plasma FGF19 levels tended to decrease with higher stages of fibrosis in NASH subjects but did not reach statistical significance. The hepatic mRNA expression (B, D and E) showed a trend for decreased expression of FXR targets

(B) SHP, **(D)** PLTP and **(E)** BSEP, while a significantly higher Cyp7A1 expression **(C)** was observed in NASH subjects ($p<0.01$).

BSEP, Bile Salt Export Pump; *Cyp7A1*, cytochrome 7α1 hydroxylase; *FGF19*, Fibroblast Growth Factor 19; *FXR*, Farsenoid-X receptor; *GAPDH*, Glyceraldehyde 3-phosphate dehydrogenase; *PLTP*, Phospholipid transfer protein; *SHP*, small heterodimer partner

Supplementary Tables 1

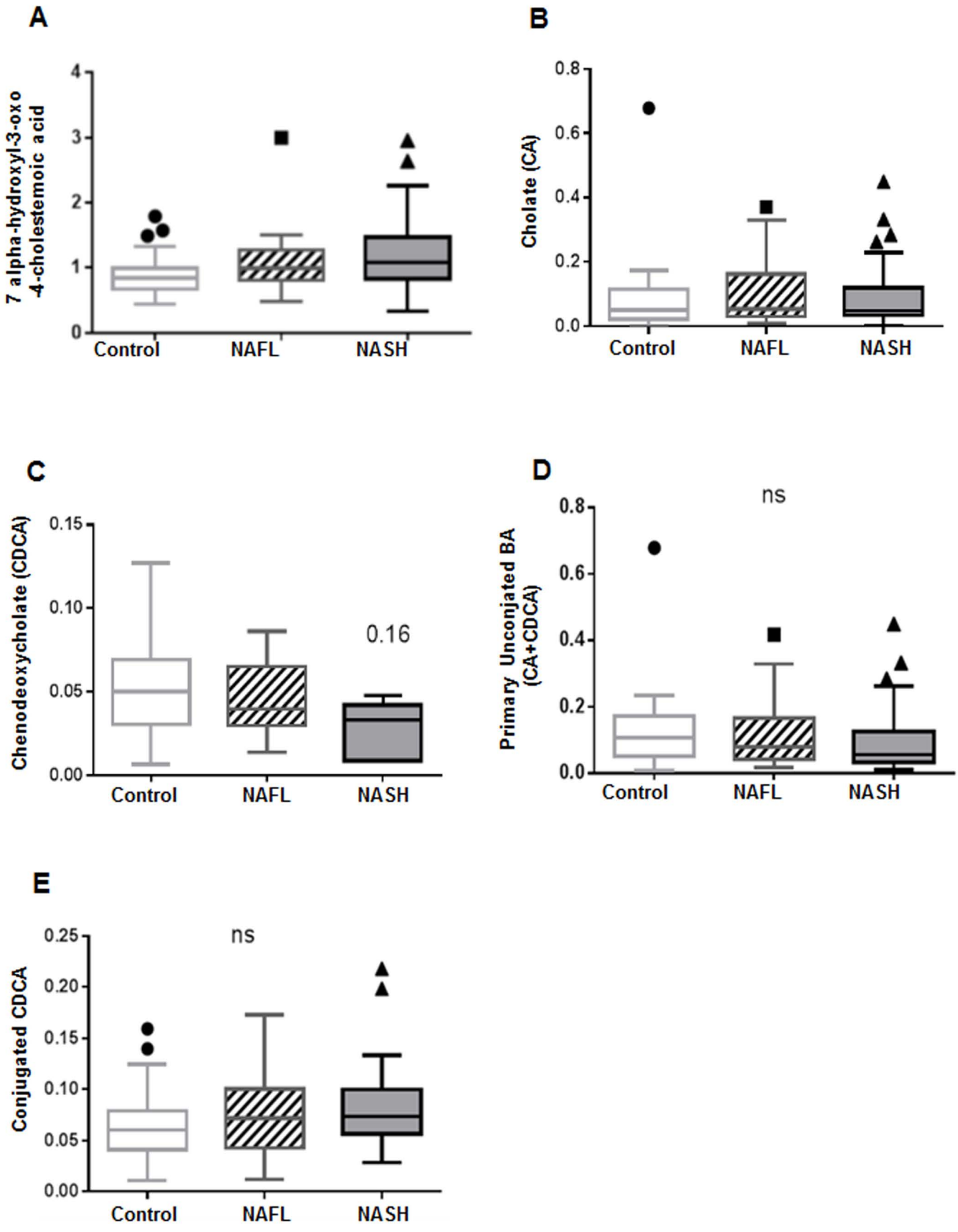
A

Parameter	ALL	Control	NAFL	NASH	p.value†
Patients - n	86	24	25	37	
Age - mean (SD)	51.8 (12.9)	39.2 (12.4)	54.6 (10.1)	58.0 (8.8)	1.32E-06
Male Sex - n (%)	29 (33.7)	11 (45.8)	10 (40.0)	8 (21.6)	0.109
BMI - mean (SD)	31.9 (5.8)	27.3 (5.8)	32.6 (5.4)	34.4 (4.2)	4.08E-05
Diabetes - n (%)	29 (33.7)	1 (4.2)	5 (20.0)	23 (62.2)	3.96E-06
Hypertension - n (%)	52 (60.5)	6 (25.0)	18 (72.0)	28 (75.7)	4.26E-05
Hyperlipidemia - n (%)	47 (54.7)	4 (16.7)	19 (76.0)	24 (64.9)	0.0469
AST - mean (SD)	37.8 (32.2)	22.3 (11.3)	45.6 (51.9)	42.4 (18.7)	7.37E-07
ALT - mean (SD)	44.1 (28.2)	22.7 (15.5)	45.5 (24.0)	57.1 (29.3)	2.75E-07
† Kruskal-Wallis tests were used to compare continuous variables and chi-square tests were used to compare categorical variables					

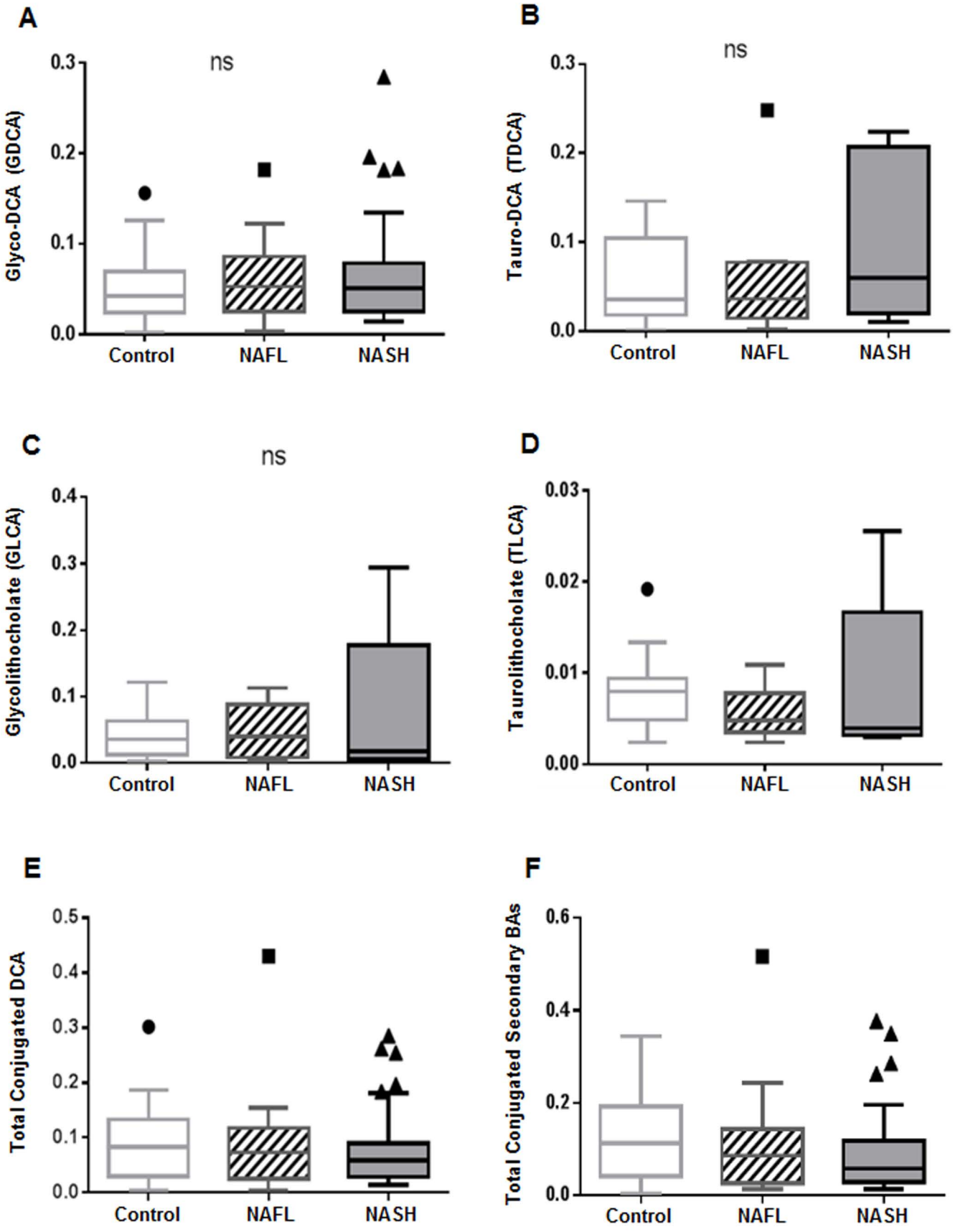
B

		F0/F1	F0/F1			
Parameter	ALL	All Patients	No NASH	NASH	F≥2	p.value†
Patients - n	83	59	44	15	24	
Age (yrs)	51.5 (13.0)	48.1 (13.1)	45.6 (13.4)	55.4 (9.0)	60.0 (8.0)	0.000156
Sex (Male) - n (%)	29 (34.9)	23 (39.0)	20 (45.5)	3 (20.0)	6 (25.0)	0.338
BMI (kg/m2)	31.8 (5.9)	31.0 (6.1)	29.5 (5.9)	35.6 (3.9)	33.7 (4.9)	0.0849
Diabetes - n (%)	29 (34.9)	13 (22.0)	4 (9.1)	9 (60.0)	16 (66.7)	0.000303
Hypertension - n (%)	50 (60.2)	32 (54.2)	20 (45.5)	12 (80.0)	18 (75.0)	0.178
Hyperlipidemia - n (%)	44 (53.0)	28 (47.5)	18 (40.9)	10 (66.7)	16 (66.7)	0.203
AST (U/L)	37.7 (32.8)	34.4 (36.3)	34.2 (41.5)	34.9 (13.3)	45.9 (20.2)	0.000213
ALT (U/L)	43.9 (28.6)	38.3 (26.8)	33.2 (23.9)	53.0 (30.4)	57.7 (28.6)	0.00155
† Mann-Whitney tests were used to compare continuous variables and chi-square tests were used to compare categorical variables. Data shown as mean (SD)						

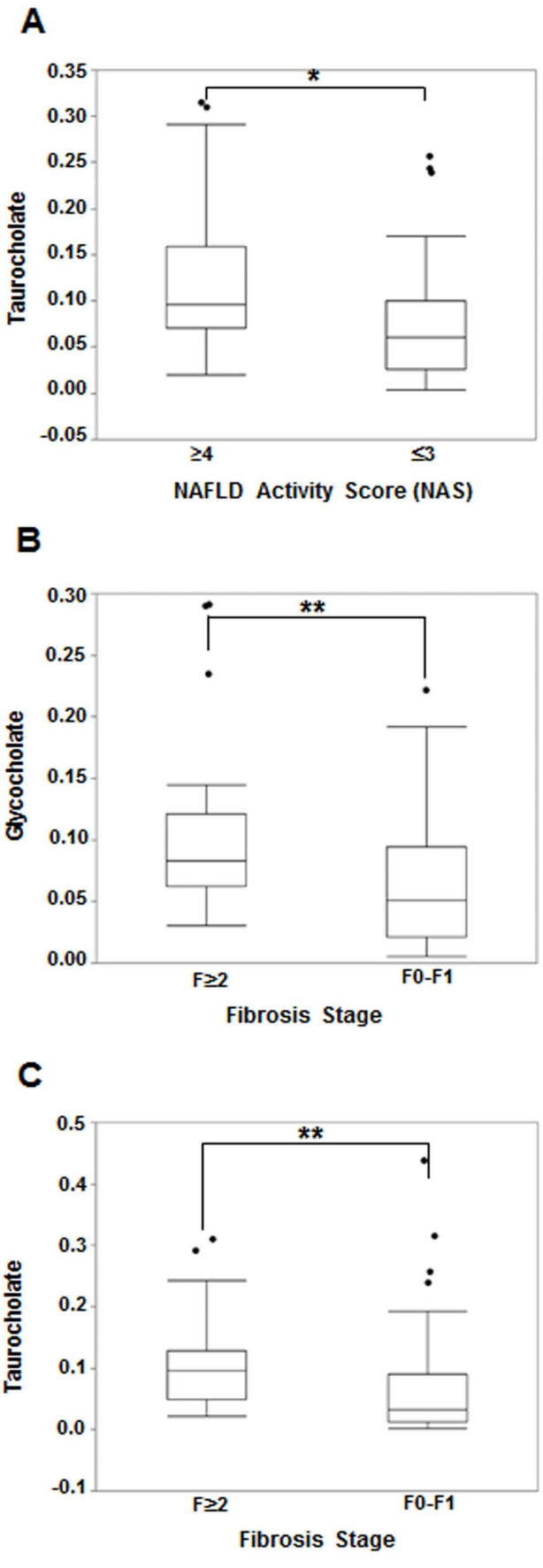
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4