

**Supplemental table 1**

Bile acid cluster	Acronym and definition
Free cholic acid	CA
Free chenodeoxycholic acid	CDCA
Free hyocholic acid	HCA
Free deoxycholic acid	DCA
Free lithocholic acid	LCA
Free ursodeoxycholic acid	UDCA
Free hyodeoxycholic acid	HDCA
Total free BAs	CA + CDCA + HCA + DCA + LCA + UDCA + HDCA
Glyco-cholic acid	GCA
Glyco-chenodeoxycholic acid	GCDCA
Glyco-hyocholic acid	GHCA
Glyco-deoxycholic acid	GDCA
Glyco-lithocholic acid	GLCA
Glyco-ursodeoxycholic acid	GUDCA
Glyco-hyodeoxycholic acid	GHDCA
Total glyco-conjugated BAs	GCA + GCDCA + GHCA + GDCA + GLCA + GUDCA + GHDCA
Tauro-cholic acid	TCA
Tauro-chenodeoxycholic acid	TCDCA
Tauro-hyocholic acid	THCA
Tauro-deoxycholic acid	TDCA
Tauro-lithocholic acid	TLCA
Tauro-ursodeoxycholic acid	TUDCA
Tauro-hyodeoxycholic acid	THDCA
Total tauro-conjugated BAs	TCA + TCDCA + THCA + TDCA + TLCA + TUDCA + THDCA
Total Conjugated BAs	Total Glyco-BAs + Total Tauro BAs
Total BAs	Total Free BAs + Total Conjugated BAs
Total CA	CA + GCA + TCA
Total CDCA	CDCA + GCDCA + TCDCA
Total HCA	HCA + GHCA + THCA
Total DCA	DCA + GDCA + THCA
Total LCA	LCA + GLCA + TLCA
Total UDCA	UDCA + GUDCA + TUDCA
Total HDCA	HDCA + GHDCA + THDCA
Total primary BAs	Total CA + Total CDCA + Total HCA
Total secondary BAs	Total DCA + Total LCA + Total UDCA + Total HDCA
Conjugated/Free BA ratio	Total Conjugated BAs / Total Free BAs
Primary/secondary BA ratio	Total primary BAs / Total secondary BAs
12 $\alpha$ -hydroxylated BAs	Total CA + Total DCA
Non-12 $\alpha$ -hydroxylated BAs	Total CDCA + Total HCA + Total LCA + Total UDCA + Total HDCA
12 $\alpha$ -hydroxylated/non-12 $\alpha$ -hydroxylated BA ratio	12 $\alpha$ -OH BAs / Non-12 $\alpha$ -OH BAs

**Supplemental Table 1.** Definition of the bile acid clusters and ratios.**Legend:** BAs: bile acids.

**Supplemental table 2**

	<b>Median (25<sup>th</sup> - 75<sup>th</sup> percentile)</b>
<b>Total BAs (nM)</b>	1612.2 (885.1-2941)
<b>Total free BAs (nM)</b>	606.9 (276.3-1262.7)
<b>Total glyco-conjugated BAs (nM)</b>	644.7 (372-1247)
<b>Total tauro-conjugated BAs (nM)</b>	84.0 (36.6-188.7)
<b>Total CA (nM)</b>	257.7 (129.9-608.4)
<b>Total CDCA (nM)</b>	648.9 (342.3-1176.2)
<b>Total DCA (nM)</b>	417.0 (208.3-777.1)
<b>Total LCA (nM)</b>	16.4 (5.1-34.8)
<b>Total UDCA (nM)</b>	81.4 (36.7-146.5)
<b>Total HCA (nM)</b>	8.4 (2.9-22)
<b>Total HDCA (nM)</b>	0 (0-0)
<b>Primary/Secondary BA ratio</b>	1.74 (1.01-3.38)
<b>Conjugated/Free BA ratio</b>	1.39 (0.57-2.94)
<b>12<math>\alpha</math>OH/non-12<math>\alpha</math>OH BA ratio</b>	1.00 (0.67-1.56)
<b>Total HCA/Total CDCA ratio</b>	0.014 (0.0063-0.0242)
<b>C4 (nM)</b>	21.8 (12.6-40.9)

**Supplemental table 2:** Baseline plasma BA concentrations and BA ratios of the studied population.

**Legend:** BA clusters are defined in the **Supplemental Table 1**.

Supplemental table 3

	Spearman's correlation coefficient					
	BMI	WHR	FPG	HbA1c	HOMA-IR	FLI
<b>CA</b>	0.02	0.08	0.11	0.06	0.09	0.04
<b>GCA</b>	0.04	0	0.04	-0.03	<b>0.17*</b>	0.08
<b>TCA</b>	0.08	-0.10	0.08	-0.01	<b>0.22**</b>	0.08
<b>CDCA</b>	0.04	0.10	0.05	0.10	0.10	0.11
<b>GCDCA</b>	-0.02	0.05	0.06	0.03	0.07	0.07
<b>TCDCA</b>	0.02	-0.12	0.08	0.06	<b>0.16*</b>	0.03
<b>DCA</b>	0.04	0.11	0.07	0	0.05	0.09
<b>GDCA</b>	-0.02	0	<b>0.14*</b>	-0.06	0.04	0.04
<b>TDCA</b>	0.05	-0.06	0.13	-0.05	0.13	0.13
<b>UDCA</b>	0.06	0.12	0.02	-0.04	0.09	0.12
<b>GUDCA</b>	-0.03	0.09	0.02	0.01	0.01	0.06
<b>TUDCA</b>	-0.04	-0.07	0.02	-0.04	0.04	0
<b>HCA</b>	-0.09	0.03	0	-0.10	-0.06	-0.05
<b>GHCA</b>	<b>-0.22**</b>	-0.07	0.04	-0.08	<b>-0.18*</b>	<b>-0.18*</b>
<b>THCA</b>	<b>-0.16*</b>	<b>-0.14*</b>	0.11	0	-0.06	<b>-0.20**</b>
<b>HDCA</b>	-0.10	-0.03	0.01	-0.08	-0.07	-0.03
<b>GHDCA</b>	-0.03	<b>0.15*</b>	0.06	-0.04	0.04	0.03
<b>THDCA</b>	0.04	0.02	0.11	0.07	0.03	0.02

**Supplemental Table 3.** Correlation between bile acids and metabolic parameters at baseline, details for free, gluco-conjugated and tauro-conjugated bile acid species.

**Legend:** Spearman's correlation coefficient is presented for each pair of parameters, and only significant p values displayed as: \*p<0.05, \*\*p<0.01 or \*\*\*p<0.001. BA clusters are defined in the Supplemental Table 1. BAs: bile acids; C4: 7α-hydroxy-4-cholest-3-one; FPG: fasting plasma glucose; G:- Gluco-conjugated; FLI: Fatty liver index; HOMA-IR: Homeostasis model assessment for insulin resistance; HR (95% CI): Hazard-Ratio with 95% confidence interval; T:- tauro-conjugated; TC: total cholesterol; TG: triglycerides; WHR: waist-hip ratio.

**Supplemental Figure 1.** Flow-chart of the selection of the patients for the bile acid study of the IT-DIAB cohort.

**Supplemental Figure 2.** Dot plots of all the significant correlations between BAs and clinic-biological parameters in Table 2.

**Supplemental Figure 3.** Correlation between HCA and C4 ( $7\alpha$ -hydroxy-4-cholesten-3-one) at baseline.

Whole population (n = 928)

n = 209 → No plasma samples (DIABENORD cohort)

Population screened (n = 719)

n = 372 → Did not fulfill inclusion criteria

Population screened (n = 347)

n = 16 → Baseline fasting plasma glucose < 110 or > 125 mg/dL

Population included (n = 331)

n = 126 → Not suitable for data analyses:  
- no follow-up visit available (n = 24)  
- no plasma for bile acids measurement (n = 7)  
- concomitant therapy with statins (n = 68) and/or fibrates (n = 11)  
- HbA<sub>1c</sub> not available (n = 1) or ≥ 6.5% (n = 16)

205 patients included in the bile acid study, of which 69 presented a conversion to new onset diabetes (33.7%) after a 5-year follow-up



