Appendix 1 Supplementary methods

Metabolite extraction, LCMS, and GCMS analyses were performed at the Swedish Metabolomics Centre (SMC, Umeå, Sweden).

Metabolite extraction

Metabolite extraction from 100 µL plasma was performed according to previously published protocols (46). A volume of 900 µL extraction buffer consisting of 90/10 v/v HPLC grade methanol (Fisher Scientific, Waltham, MA, USA): milliQ containing internal standards were added followed by shaking at 30 Hz for 2 minutes in a mixer mill. Proteins were precipitated on ice at 4 °C, followed by centrifugation at 4 °C, 14,000 rpm, and 10 minutes. The supernatant, 100 µL for GCMS and 200 µL for LCMS, was added to microvials and evaporated to dryness using a speed-vac concentrator. Solvents were then evaporated and samples placed in a -80 °C freezer until analysis. Quality control samples, consisting of a pool of small amounts of leftover supernatants, were analyzed by MSMS (LCMS) to identify metabolites.

LCMS analysis

Samples were mixed with 10 + 10 µL methanol and water and run in both negative and positive mode. Chromatography was performed on Agilent 1290 Infinity UHPLC-system (Agilent Technologies, Waldbronn, Germany) and subsequent detection using Agilent 6550 Q-TOF mass spectrometer linked with a jet stream electrospray ion source. Agilent Masshunter Profinder version B.08.00 (Agilent Technologies Inc., Santa Clara, CA, USA) was used to process data. A pre-specified list of commonly detected metabolites in serum or plasma was searched for using Batch Targeted feature extraction in Masshunter Profinder. An inhouse LCMS library generated using authentic standards was used for targeted processing. Information on MS, MSMS, and retention time was used to identify metabolites.

GCMS analysis

Derivatization and GC-MS analysis was run on a Pegasus HT time-of-flight mass spectrometer, GC/TOFMS (Leco Corp., St Joseph, MI) as described previously (46). Non-processed MS-files extracted from ChromaTOF software were imported to MATLAB® 2018a (Mathworks, Natick, MA, USA) where the following steps were taken; base-line correction, chromatogram alignment, data compression, and multivariate curve resolution (47). Mass spectra were recognized by comparing retention index and mass spectra of available metabolites in libraries performed using the NIST MS 2.2. software (48). Both reverse and forward searches were performed, and more carefully so on masses and ratio between masses suggestive of a derivatized metabolite. A peak was annotated as the metabolite with the mass spectrum with the highest probability and a maximum difference of five between sample and library for the suggested metabolite.

Imputation

Some participants lacked information on smoking and BMI. BMI was missing for 3 % (n=5) and smoking status for 0.6 % (n=1) of the study cohort. The reason for missingness is unknown. Multiple imputation chained equation (mice) was performed to impute missing values in the pre-diagnostic cohort using mice R package version 3.14.0 (21). BMI and smoking were imputed six times. Included variables in the imputation for logistic regression models were fasting status, case-control set, BMI, age, sex, smoking status, and case/control status (outcome variable). The adjusted logistic regression model was performed for each imputed dataset and the results pooled using Rubin's rules (average). Predictive mean matching (pmm) was chosen as imputation method. For the sub-cohort based on symptoms, only one individual missed information on smoking and thus the mean value of smoking status in the cohort was imputed. For survival analysis, BMI was imputed by the mean value. For the final logistic regression models of LASSO selected variables, BMI and smoking were imputed with mean values.

https://dx.doi.org/10.21037/jgo-23-930

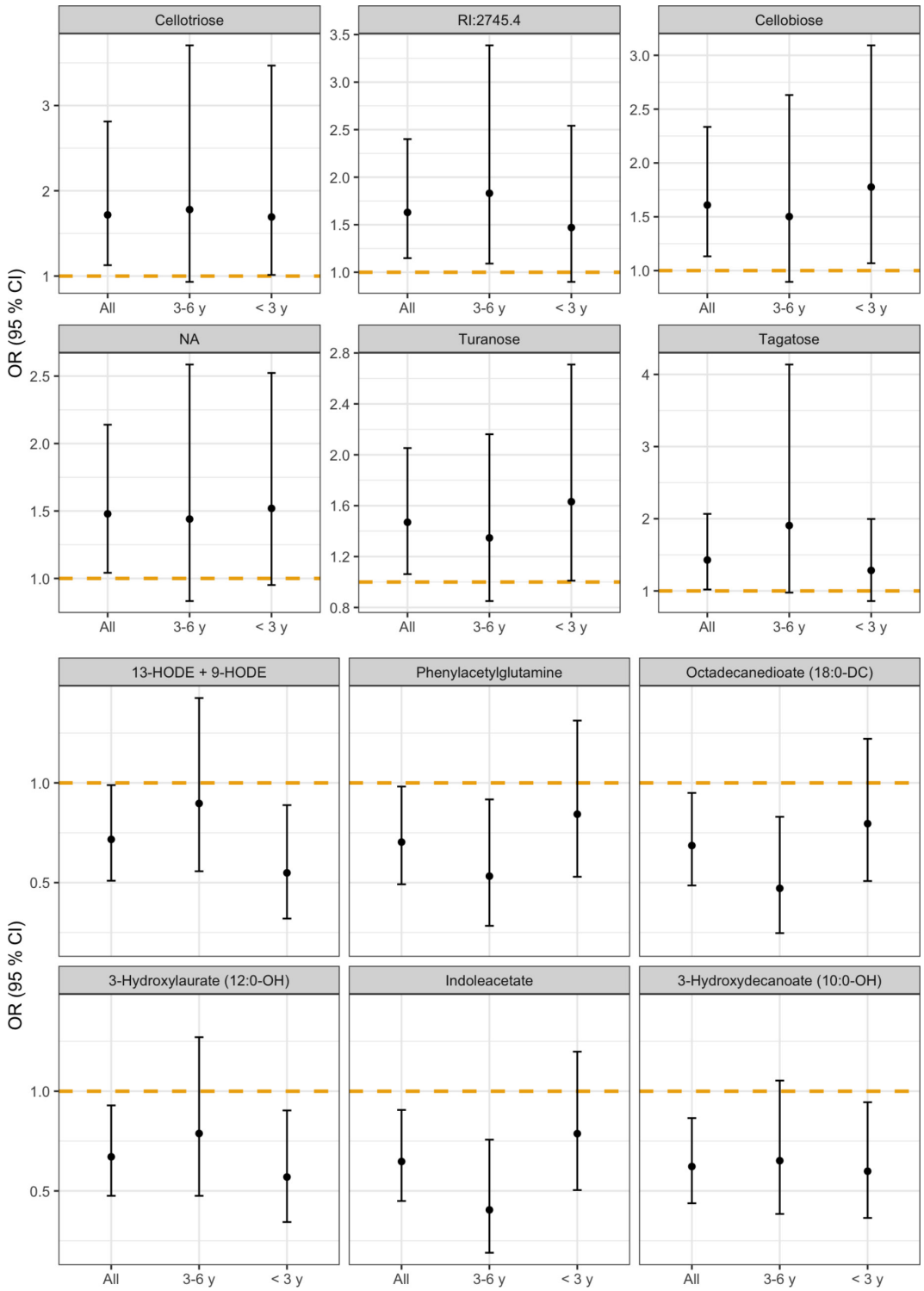

Figure S1 Plasma metabolites with a nominal P value <0.05 in the pre-diagnostic cohort ('All' = 82 future pancreatic cancer cases and 82 matched healthy controls), 3-6 years (y) lag-time, or <3 y lag-time to PDAC diagnosis. Odds ratios (OR) as well as 95% confidence intervals (CI) are shown. Horizontal dashed yellow line represents an OR of one.

https://dx.doi.org/10.21037/jgo-23-930

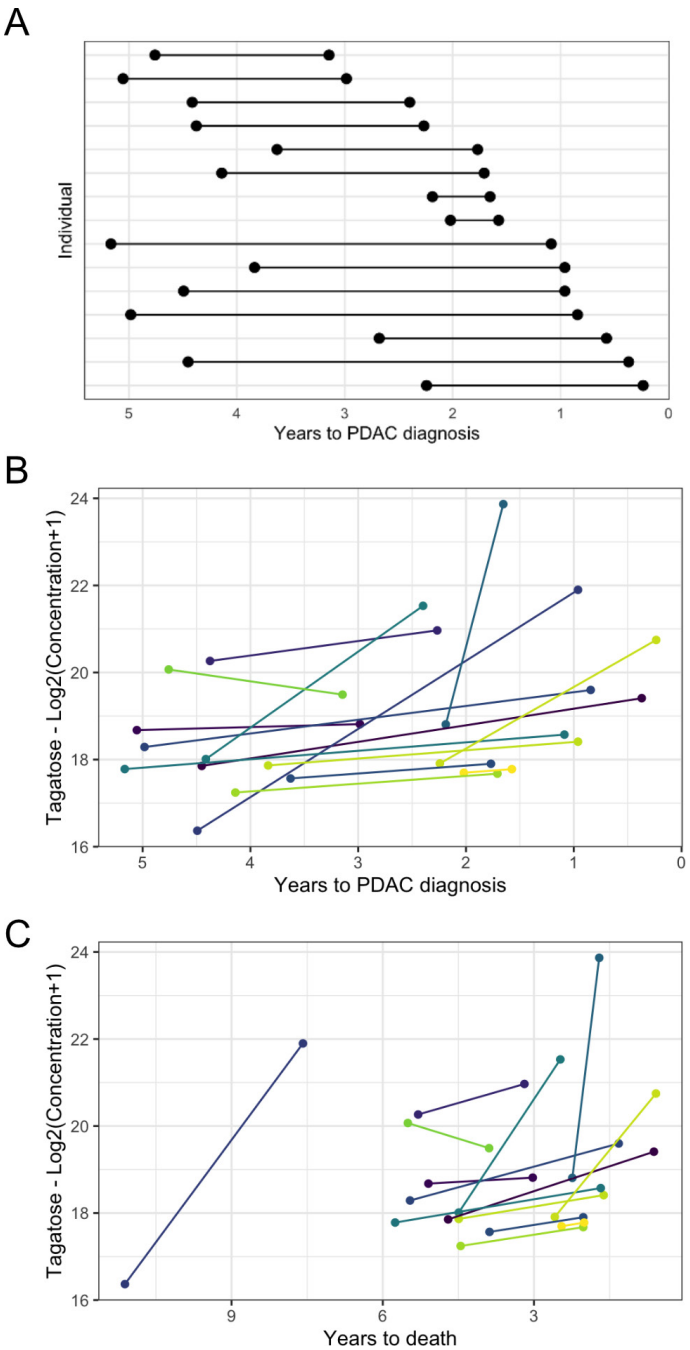

Figure S2 Longitudinal analysis of D-tagatose in future pancreatic cancer patients (n=14). Time between sampling and pancreatic cancer diagnosis for each individual is shown for the two sampling occasions (A). Tagatose levels in relation to time to pancreatic cancer diagnosis (B) and time to death (C). Each line represents one patient.

https://dx.doi.org/10.21037/jgo-23-930

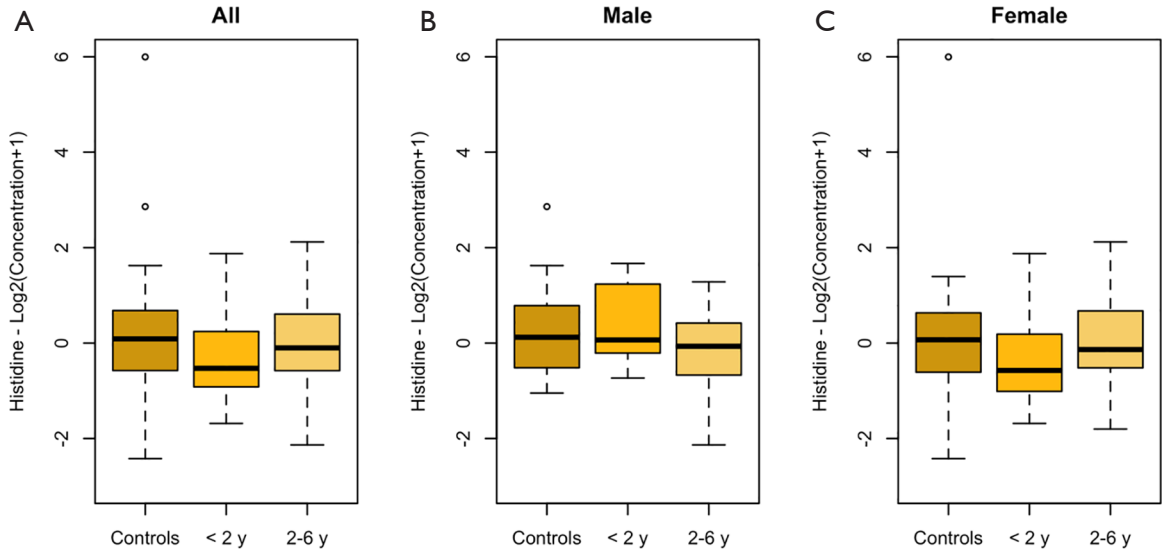

Figure S3 Circulating histidine levels in (A) all individuals, (B) only males or (C) only females.

https://dx.doi.org/10.21037/jgo-23-930

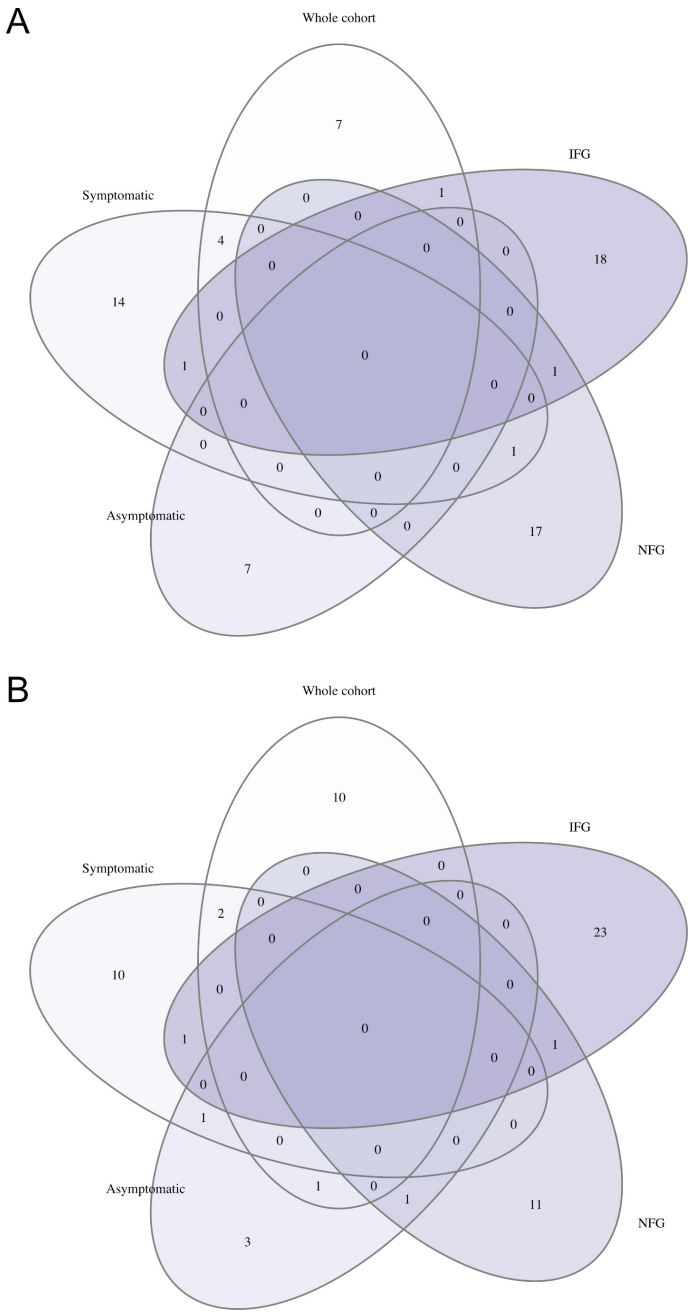

Figure S4 Overlap of metabolite profiles in 'Crude' models (A) and 'Adjusted' models (B). IFG, impaired fasting glucose; NFG, normal fasting glucose.

https://dx.doi.org/10.21037/jgo-23-930

Table S1 Inclusion and exclusion criteria of pre-diagnostic symptoms

Prediagnostic symptom	Inclusion	Exclusion
Jaundice	Stated jaundice; yellow sclerae or skin	
Newly diagnosed diabetes	A diabetes mellitus diagnosis ≤3 years prior to pancreatic cancer diagnosis	Elevated blood glucose elevations only A diabetes mellitus diagnosis >3 years before pancreatic cancer diagnosis
Weight loss	Described weight loss	
Back pain	General back pain; lumbago; thoracic spine back pain; pain radiating towards the back	Pain restricted to cervical pain
Abdominal pain	General abdominal pain; burning, stinging, dull pain sensation; abdominal discomfort	Diffuse abdominal problems
Diarrhea	Diarrhea stated	
Fatigue	General fatigue	
Gallstone disease	Gallstone diagnosis	
Pancreatitis	Pancreatitis diagnosis	

Table S2 Logistic regression models of circulating metabolites in pre-diagnostic pancreatic cancer (82 cases, 82 controls)

Metabolite	Crude OR (95% CI)	Crude P value	Adjusted OR (95% CI)	Adjusted P value	Method
13-HODE+9-HODE	0.72 (0.51-0.99)	0.048	0.68 (0.47-0.99)	0.048	LCMS
3-Hydroxydecanoate (10:0-OH)[1]	0.62 (0.44-0.87)	0.006	0.61 (0.41-0.89)	0.012	LCMS
3-Hydroxylaurate (12:0-OH)	0.67 (0.48-0.93)	0.019	0.64 (0.43-0.95)	0.028	LCMS
Cellobiose	1.61 (1.13-2.33)	0.01	1.88 (1.25-2.84)	0.003	GCMS
Cellotriose	1.72 (1.13-2.81)	0.023	1.86 (1.13-3.06)	0.017	GCMS
D-Tagatose	1.43 (1.02-2.07)	0.046	1.61 (1.09-2.38)	0.019	GCMS
Indoleacetate[1]	0.65 (0.45-0.91)	0.014	0.64 (0.45-0.91)	0.014	LCMS
Octadecanedioate (18:0-DC)	0.69 (0.49-0.95)	0.027	0.75 (0.52-1.06)	0.105	LCMS
Phenylacetylglutamine	0.7 (0.49-0.98)	0.044	0.74 (0.52-1.06)	0.104	LCMS
RI: 2745.4[1]	1.63 (1.15-2.4)	0.009	1.67 (1.13-2.49)	0.012	GCMS
Turanose	1.47 (1.06-2.05)	0.022	1.53 (1.08-2.16)	0.017	GCMS
Maltose	1.48 (1.04-2.14)	0.032	1.65 (1.1-2.45)	0.015	GCMS

[1], selected by LASSO in ≥70% of random bootstrapping subsets. Crude models were adjusted for matching factors; age, sex and storage time. Models were further adjusted for body mass index, fasting status, and smoking status ('Adjusted'). Smoking and BMI were imputed for one and three individuals, respectively. Nominal P values were derived from Wald test of the metabolite coefficient in logistic regression models. No metabolite remained significant after adjusting for multiple hypothesis testing using Benjamini-Hochberg. OR, odds ratio per standard deviation increase; CI, confidence interval; GCMS, gas chromatography mass spectrometry; LCMS, liquid chromatography mass spectrometry; RI, retention index.

https://dx.doi.org/10.21037/jgo-23-930

Table S3 Conditional logistic regression models of circulating metabolites in pre-diagnostic pancreatic cancer (82 cases, 82 controls)

Metabolite	Crude OR (95% CI)	Crude P value	Adjusted OR (95% CI)	Adjusted P value	Method
1-oxo-stearoyl-GPC (16-0-ox)	0.51 (0.29-0.91)	0.023	0.39 (0.16-0.93)	0.037	LCMS
13-HODE + 9-HODE	0.58 (0.38-0.88)	0.011	0.36 (0.17-0.76)	0.009	LCMS
3-Hydroxydecanoate (10:0-OH)	0.53 (0.36-0.79)	0.002	0.5 (0.28-0.89)	0.020	LCMS
3-Hydroxylaurate (12:0-OH)	0.56 (0.37-0.85)	0.006	0.48 (0.26-0.87)	0.019	LCMS
Arabinose	3.12 (1.25-7.76)	0.015	3.47 (0.98-12.32)	0.058	GCMS
Cellobiose	1.75 (1.13-2.71)	0.012	1.78 (0.97-3.26)	0.066	GCMS
Cellotriose	1.76 (1.06-2.91)	0.029	2.96 (1.21-7.22)	0.020	GCMS
Creatinine	1.82 (1.11-2.98)	0.018	1.91 (0.97-3.76)	0.064	GCMS
D-Tagatose	1.6 (1.03-2.49)	0.038	3.02 (1.11-8.21)	0.033	GCMS
Dimethylarginine (ADMA + SDMA)	0.66 (0.45-0.96)	0.031	0.6 (0.34-1.05)	0.078	LCMS
DL-beta-Hydroxybutyric acid 1	0.47 (0.24-0.91)	0.026	0.7 (0.35-1.38)	0.303	GCMS
Hydroxylauroyl-carnitine (C12:0-OH)	0.47 (0.25-0.86)	0.015	0.35 (0.13-0.91)	0.035	LCMS
Hydroxymyristate (14:0-OH)	0.58 (0.38-0.89)	0.013	0.49 (0.26-0.93)	0.034	LCMS
Hydroxyoctadecenoyl-carnitine (C18:1-OH)	0.65 (0.43-0.97)	0.036	0.57 (0.3-1.06)	0.081	LCMS
Hydroxystearate (18:0-OH)	0.66 (0.45-0.96)	0.030	0.46 (0.24-0.87)	0.020	LCMS
Indoleacetate	0.69 (0.49-0.97)	0.035	0.57 (0.35-0.92)	0.025	LCMS
Octadecanedioate (18:0-DC)	0.62 (0.43-0.9)	0.012	0.6 (0.35-1.05)	0.076	LCMS
Phenylacetylglutamine	0.66 (0.45-0.97)	0.034	0.71 (0.44-1.13)	0.150	LCMS
Phenylalanylalanine	0.64 (0.41-0.99)	0.045	0.63 (0.34-1.15)	0.138	LCMS
RI: 1569.3	1.58 (1.05-2.39)	0.029	1.38 (0.74-2.58)	0.314	GCMS
RI: 2745.4	1.7 (1.11-2.6)	0.014	1.42 (0.79-2.58)	0.246	GCMS
Turanose	1.6 (1.09-2.37)	0.018	1.38 (0.8-2.37)	0.252	GCMS

Models with or without adjustment for storage time, body mass index (BMI), fasting status, and smoking ('Adjusted') are shown. Smoking and BMI were imputed for three and ten individuals, respectively. Nominal P values were derived from Wald test of the metabolite coefficient in conditional logistic regression models. No metabolite remained significant after adjusting for multiple hypothesis testing using Benjamini-Hochberg. OR, odds ratio per standard deviation increase; CI, confidence interval; GCMS, gas chromatography mass spectrometry; LCMS, liquid chromatography mass spectrometry; RI, retention index.

https://dx.doi.org/10.21037/jgo-23-930

Table S4 Reported PDAC symptoms up to six years prior to diagnosis. Medical records were available for 76 patients

Symptom	Number of patients	Percentage of patients (%)
Abdominal pain	64	84
Weight loss	37	48
Back pain	37	48
Fatigue	36	47
Jaundice	26	34
Diarrhea	21	27
Newly diagnosed diabetes	10	13
Gallstone disease	5	6
Pancreatitis	0	0

Table S5 Conditional logistic regression models of circulating metabolites in the symptomatic cohort (34 future pancreatic cancer cases and 34 matched healthy controls)

Metabolite	Crude OR (95% CI)	Crude P value	Adjusted OR (95% CI)	Adjusted P value	Method
1-Eicosapentaenoyl-GPE (20:5)	1.82 (1.01-3.28)	0.05	2.21 (1.02-4.79)	0.04	LCMS
2-Eicosatrienoyl-GPC (20:3)	1.8 (1.01-3.2)	0.04	1.5 (0.82-2.75)	0.19	LCMS
2-Palmitoleoyl-GPC (16:1)	2.89 (1.27-6.58)	0.01	2.45 (1.02-5.87)	0.04	LCMS
3-Hydroxydecanoate (10:0-OH)	0.52 (0.28-0.96)	0.04	0.58 (0.32-1.09)	0.09	LCMS
3-Hydroxylaurate (12:0-OH)	0.52 (0.27-0.99)	0.05	0.58 (0.3-1.12)	0.11	LCMS
5-Hydroxyhexanoate (6:0-OH)	1.99 (1.01-3.94)	0.05	2 (0.87-4.59)	0.10	LCMS
Acetaminophen (drug)	1.74 (1-3.02)	0.05	1.97 (1.03-3.78)	0.04	GCMS
Chenodeoxycholic acid glycine conjugate	2.31 (1.17-4.58)	0.02	2.27 (0.98-5.27)	0.06	LCMS
Creatinine	2.87 (1.1-7.51)	0.03	2.83 (0.99-8.1)	0.05	GCMS
Hyocholic acid glycine conjugate	1.88 (1.04-3.39)	0.04	1.69 (0.89-3.2)	0.11	LCMS
Indoleacetate	0.53 (0.29-0.98)	0.04	0.48 (0.23-0.98)	0.04	LCMS
Phenylalanine	2.33 (1.11-4.86)	0.02	2.51 (1.11-5.67)	0.03	GCMS
Phenylalanine	2.15 (1.08-4.3)	0.03	2.19 (1.06-4.52)	0.03	GCMS
Pyroglutamic acid	2.02 (1.05-3.9)	0.04	2.12 (1.02-4.39)	0.04	GCMS
RI: 1427.5	2.11 (1.01-4.43)	0.05	1.81 (0.79-4.14)	0.16	GCMS
RI: 1516	2.02 (1.04-3.93)	0.04	2.11 (1.01-4.41)	0.05	GCMS
RI: 1569.3	2.07 (1.01-4.26)	0.05	2 (0.89-4.49)	0.09	GCMS
Taurochenodesoxycholic acid	2.33 (1.17-4.62)	0.02	1.99 (0.93-4.28)	0.08	LCMS
Taurocholcholic acid	1.92 (1.04-3.56)	0.04	1.54 (0.78-3.05)	0.21	LCMS
Turanose	2.25 (1.07-4.74)	0.03	2.65 (1.12-6.27)	0.03	GCMS

Models were adjusted for body mass index (BMI), fasting status, and smoking. Nominal P values of the metabolite coefficient are shown, no metabolite remained significant after adjusting for multiple hypothesis testing using Benjamini-Hochberg. OR, odds ratio per standard deviation increase; CI, confidence interval; GCMS, gas chromatography mass spectrometry; LCMS, liquid chromatography mass spectrometry; RI, retention index.

https://dx.doi.org/10.21037/jgo-23-930

Table S6 Conditional logistic regression models of circulating metabolites in the asymptomatic cohort (44 future pancreatic cancer cases and 44 matched healthy controls)

Metabolite	Crude OR (95% CI)	Crude P value	Adjusted OR (95% CI)	Adjusted P value	Method
Arabinose	2.37 (1.05-5.37)	0.04	3.48 (1.26-9.56)	0.02	GCMS
Cholic acid	0.52 (0.3-0.91)	0.02	0.56 (0.31-0.98)	0.05	LCMS
Erythronic acid	0.57 (0.33-0.96)	0.03	0.58 (0.34-0.98)	0.05	GCMS
Hydroxylauroyl-carnitine (C12:0-OH)	0.42 (0.18-1)	0.05	0.42 (0.17-1.05)	0.07	LCMS
Hydroxyoctadecenoyl-carnitine (C18:1-OH)	0.52 (0.27-0.99)	0.05	0.5 (0.25-1.01)	0.06	LCMS
Linoleneoyl-carnitine (C18:3)	0.52 (0.28-0.99)	0.05	0.53 (0.27-1.03)	0.07	LCMS
Palmitoylcarnitine (C16:0)	0.5 (0.28-0.88)	0.02	0.44 (0.22-0.87)	0.02	LCMS

Models were adjusted for body mass index (BMI), fasting status, and smoking. Nominal P values of the metabolite coefficient are shown, no metabolite remained significant after adjusting for multiple hypothesis testing using Benjamini-Hochberg. OR, odds ratio per standard deviation increase; CI, confidence interval; GCMS, gas chromatography mass spectrometry; LCMS, liquid chromatography mass spectrometry.

Table S7 Altered metabolites between future pancreatic cancer patients with or without symptoms

Metabolite	Crude OR (95% CI)	Crude P value	Adjusted OR (95% CI)	Adjusted P value	Method
5-Hydroxyhexanoate (6:0-OH)	1.72 (1.07-2.95)	0.03	1.74 (0.98-3.29)	0.07	LCMS
Chenodeoxycholic acid glycine conjugate	1.95 (1.19-3.38)	0.01	2.32 (1.21-4.97)	0.02	LCMS
Glutamic acid	1.79 (1.1-3.2)	0.03	1.63 (0.92-3.09)	0.11	GCMS
Glycoursodeoxycholic acid	1.73 (1.08-2.95)	0.03	1.75 (1.02-3.15)	0.05	LCMS
Hyocholic acid glycine conjugate	2.01 (1.23-3.54)	0.01	2.17 (1.21-4.26)	0.01	LCMS
Isoleucine	1.69 (1.06-2.83)	0.03	1.58 (0.93-2.8)	0.10	GCMS
Leucine	1.77 (1.09-3.08)	0.03	1.67 (0.98-3.03)	0.07	GCMS
Lithocholic acid glycine conjugate	1.68 (1.06-2.79)	0.03	1.75 (1-3.23)	0.06	LCMS
Phenylalanine	1.68 (1.05-2.83)	0.04	1.55 (0.9-2.78)	0.12	GCMS
Pyroglutamic acid	1.67 (1.04-2.85)	0.04	1.72 (0.99-3.16)	0.06	GCMS
RI: 1516	1.66 (1.03-2.83)	0.05	1.7 (0.99-3.13)	0.07	GCMS
RI: 2350	0.58 (0.31-0.95)	0.05	0.53 (0.27-0.91)	0.03	GCMS
Taurochenodesoxycholic acid	1.67 (1.05-2.8)	0.04	1.81 (1.04-3.32)	0.04	LCMS
Taurocholcholic acid	1.7 (1.06-2.87)	0.04	1.86 (1.06-3.49)	0.04	LCMS
Tryptophan	1.8 (1.09-3.21)	0.03	1.85 (1.06-3.53)	0.04	GCMS

Nominally significant metabolites from unconditional logistic regression of future pancreatic cancer patients with symptoms (n=34) versus without symptoms (n=44). Models were adjusted for age, sex, sample date, BMI, smoking status, and fasting status ('Adjusted'). Nominal P values are shown and no metabolite remained significant after adjusting for multiple hypothesis testing. OR, odds ratio per standard deviation increase; CI, confidence interval; LCMS, liquid chromatography mass spectrometry; GCMS, gas chromatography mass spectrometry; RI, retention index.

https://dx.doi.org/10.21037/jgo-23-930

Table S8 Altered metabolites between future pancreatic cancer patients and healthy controls with impaired fasting glucose

Metabolite	Crude OR (95% CI)	Crude P value	Adjusted OR (95% CI)	Adjusted P value	Method
1-palmitoyl-plasmanylethanolamine	2.95 (1.1-7.89)	0.031	3.99 (1.18-13.5)	0.026	LCMS
5-hydroxyindoleacetate	0.39 (0.16-0.97)	0.044	0.24 (0.07-0.78)	0.018	LCMS
Acisoga	2.91 (1.05-8.02)	0.04	2.98 (0.98-9.14)	0.055	LCMS
Arabinose	0.13 (0.03-0.64)	0.012	0.08 (0.01-0.63)	0.017	GCMS
Arabitol	0.28 (0.1-0.73)	0.009	0.19 (0.05-0.71)	0.013	GCMS
Citrulline or Arginine	3.04 (1.08-8.52)	0.035	3.96 (1.26-12.39)	0.018	GCMS
Dodecanoic acid	0.39 (0.16-0.93)	0.033	0.39 (0.16-0.99)	0.047	GCMS
Hypoxanthine	0.26 (0.09-0.77)	0.015	0.24 (0.07-0.76)	0.015	LCMS
Kynurenine	0.37 (0.15-0.94)	0.036	0.2 (0.06-0.7)	0.012	LCMS
N6-succinyladenosine	0.3 (0.1-0.85)	0.024	0.21 (0.06-0.75)	0.016	LCMS
Octadecanedioate (18:0-DC)	0.34 (0.12-0.96)	0.041	0.36 (0.13-1.02)	0.055	LCMS
Pentose	0.34 (0.12-0.98)	0.045	0.28 (0.08-0.98)	0.047	GCMS
Phenylalanine	0.26 (0.09-0.79)	0.017	0.26 (0.09-0.82)	0.022	GCMS
RI: 1427.5	0.25 (0.07-0.92)	0.036	0.21 (0.06-0.83)	0.025	GCMS
RI: 1566.1	0.02 (0-0.94)	0.046	0.01 (0-0.95)	0.048	GCMS
RI: 1622.5	2.53 (1.01-6.38)	0.048	2.85 (1.08-7.53)	0.034	GCMS
RI: 1633.2 (Fatty acid)	0.38 (0.16-0.93)	0.034	0.38 (0.15-0.98)	0.045	GCMS
Ribitol	0.28 (0.1-0.8)	0.018	0.21 (0.05-0.77)	0.019	GCMS
Xylose[1]	0.17 (0.04-0.77)	0.022	0.02 (0-0.51)	0.017	GCMS

[1], five future PDAC outliers with a value of 0 was excluded from statistical analysis. Logistic regression models in individuals with impaired fasting glucose levels (19 future PDAC cases, 15 healthy controls) without ('Crude') and with adjustment for age, sex, BMI, smoking status and sample date ('Adjusted'). No metabolite remained significant after adjusting for multiple hypothesis testing using Benjamini-Hochberg method. OR, odds ratio per standard deviation increase; CI, confidence interval; GCMS, gas chromatography mass spectrometry; LCMS, liquid chromatography mass spectrometry; RI, retention index.

https://dx.doi.org/10.21037/jgo-23-930

Table S9 Altered metabolites between future pancreatic cancer patients and healthy controls with normal fasting glucose

Metabolite	Crude OR (95% CI)	Crude P value	Adjusted OR (95% CI)	Adjusted P value	Method
1-oleoyl-GPC (18:1)	2.17 (1.09-4.32)	0.027	2.31 (1.11-4.8)	0.025	LCMS
1-palmitoyl-plasmanyl-GPC (C16:0)	0.59 (0.35-0.99)	0.047	0.62 (0.37-1.05)	0.077	LCMS
2-linoleoyl-GPC (18:2)	1.88 (1.11-3.19)	0.02	2.1 (1.15-3.83)	0.016	LCMS
D-Tagatose	0.60 (0.36-0.99)	0.044	0.55 (0.32-0.94)	0.028	GCMS
Galacturonic acid	2.26 (1.23-4.16)	0.008	2.67 (1.33-5.37)	0.006	GCMS
Glucose	1.73 (1.02-2.94)	0.044	1.67 (0.96-2.93)	0.07	GCMS
Glucuronic acid	1.73 (1.01-2.96)	0.046	1.76 (1.01-3.05)	0.045	GCMS
Gulose	1.74 (1.00-3.00)	0.048	1.67 (0.94-2.98)	0.081	GCMS
Homocitrulline	0.56 (0.33-0.94)	0.029	0.56 (0.32-0.96)	0.036	LCMS
Phenylalanine	1.74 (1.02-2.96)	0.041	1.74 (1.00-3.02)	0.051	GCMS
Phosphoric acid	2.26 (1.19-4.26)	0.012	2.19 (1.11-4.32)	0.025	GCMS
RI: 1895.3 (sugar)	1.88 (1.01-3.53)	0.048	1.8 (0.93-3.49)	0.079	GCMS
RI: 1913 (sugar)	1.74 (1.01-3.00)	0.047	1.7 (0.95-3.03)	0.073	GCMS
RI: 2350.4	0.56 (0.32-0.99)	0.048	0.52 (0.27-0.99)	0.047	GCMS
RI: 2781.5	1.82 (1.09-3.05)	0.022	1.81 (1.08-3.06)	0.025	GCMS
RI: 3501.6	1.71 (1.02-2.89)	0.043	1.65 (0.97-2.83)	0.067	GCMS
Xylose	1.68 (1.01-2.80)	0.046	1.72 (1.00-2.96)	0.052	GCMS

Logistic regression models in individuals with normal fasting glucose levels (36 future PDAC cases, 40 healthy controls) without ('Crude') and with adjustment for age, sex, BMI, smoking status and sample date ('Adjusted'). No metabolite remained significant after adjusting for multiple hypothesis testing using Benjamini-Hochberg method. OR, odds ratio per standard deviation increase; CI, confidence interval; GCMS, gas chromatography mass spectrometry; LCMS, liquid chromatography mass spectrometry; RI, retention index.

https://dx.doi.org/10.21037/jgo-23-930

Table S10 Altered metabolites between future pancreatic cancer patients with normal or impaired fasting glucose levels

Metabolite	Crude OR (95% CI)	Crude P value	Adjusted OR (95% CI)	Adjusted P value	Method
1-linolenoyl-GPC (18:3n3)	0.51 (0.25-0.93)	0.037	0.48 (0.23-0.94)	0.042	LCMS
1-linoleoyl-GPC (18:2)	0.54 (0.28-0.97)	0.050	0.45 (0.21-0.87)	0.024	LCMS
2-linoleoyl-GPC (18:2)	0.43 (0.2-0.82)	0.018	0.38 (0.15-0.79)	0.017	LCMS
Alpha-Tocopherol	0.48 (0.22-0.91)	0.043	0.43 (0.17-0.87)	0.034	GCMS
Arabinose	0.51 (0.26-0.93)	0.038	0.53 (0.25-1.02)	0.070	GCMS
Arabitol	0.23 (0.08-0.52)	0.002	0.19 (0.06-0.46)	0.001	GCMS
Glycerol-2-phosphate	0.41 (0.16-0.85)	0.035	0.38 (0.13-0.91)	0.050	GCMS
Hexadeanoic acid (IS)	0.42 (0.15-0.87)	0.046	0.39 (0.14-0.86)	0.038	GCMS
Hexadecenoyl carnitine (C16:1)	2.05 (1.11-4.22)	0.031	1.93 (0.99-4.18)	0.068	LCMS
Isoleucylisoleucine	0.4 (0.19-0.76)	0.009	0.43 (0.19-0.88)	0.028	LCMS
Myristoleate (14:1)	2.04 (1.08-4.35)	0.042	1.91 (0.9-4.75)	0.121	LCMS
Oleate (18:1)	1.88 (1.04-3.67)	0.046	1.86 (0.93-4.2)	0.099	LCMS
Oleoylcarnitine (C18:1)	1.96 (1.07-3.98)	0.042	2.04 (1.07-4.31)	0.042	LCMS
Palmitoleate (16:1)	2.45 (1.27-5.39)	0.014	2.27 (1.08-5.71)	0.049	LCMS
Phenylalanylleucine	0.4 (0.17-0.82)	0.020	0.41 (0.17-0.83)	0.028	LCMS
Phenylalanylphenylalanine	0.54 (0.28-0.96)	0.044	0.58 (0.29-1.11)	0.109	LCMS
Phosphoric acid	0.27 (0.09-0.65)	0.010	0.23 (0.07-0.59)	0.006	GCMS
RI: 1408	0.37 (0.14-0.8)	0.027	0.31 (0.1-0.76)	0.027	GCMS
RI: 1427.5	0.43 (0.2-0.85)	0.025	0.43 (0.18-0.91)	0.039	GCMS
RI: 1616	0.07 (0-0.64)	0.044	0.02 (0-0.34)	0.016	GCMS
RI: 1653	0.53 (0.27-0.95)	0.042	0.48 (0.22-0.94)	0.045	GCMS
RI: 2781.5	0.41 (0.19-0.79)	0.014	0.4 (0.16-0.84)	0.028	GCMS
trans-4-Hydroxy-L-proline	0.45 (0.2-0.89)	0.037	0.49 (0.2-1.07)	0.092	GCMS
Xylose	0.46 (0.21-0.88)	0.031	0.43 (0.17-0.88)	0.039	GCMS

Nominally significant metabolites from unconditional logistic regression of future pancreatic cancer patients with impaired fasting glucose (n=19) versus future pancreatic cancer patients with normal fasting glucose (n=36). Models were adjusted for age, sex, and sample date, BMI, and smoking status. Nominal P values are shown, no metabolite remained significant after adjusting for multiple hypothesis testing. OR, odds ratio per standard deviation increase; LCMS, liquid chromatography mass spectrometry; GCMS, gas chromatography mass spectrometry; RI, retention index; IS, internal standard.

https://dx.doi.org/10.21037/jgo-23-930

Table S11 Cox regression analysis in pre-diagnostic pancreatic cancer

Metabolite	Crude HR (95% CI)	Crude P value	Crude FDR	Adjusted HR (95% CI)	Adjusted P value	Adjusted FDR
13-HODE + 9-HODE	2.77 (1.62-4.71)	<0.001	0.01	3.04 (1.67-5.54)	<0.001	0.01
2-Hydroxyhexanoate (6:0-OH)	2.91 (1.60-5.27)	<0.001	0.02	2.84 (1.48-5.43)	0.002	0.02
2-Hydroxyoctanoate (8:0-OH)	2.31 (1.29-4.14)	0.005	0.08	2.94 (1.56-5.55)	0.001	0.02
3-Hydroxydecanoate (10:0-OH)	2.79 (1.61-4.81)	<0.001	0.01	2.78 (1.6-4.83)	<0.001	0.01
3-Hydroxylaurate (12:0-OH)	2.61 (1.50-4.54)	0.001	0.02	2.61 (1.41-4.82)	0.002	0.02
3-Hydroxyoctanoate (8:0-OH)	1.90 (1.22-2.97)	0.004	0.08	1.98 (1.27-3.07)	0.002	0.02
3-Hydroxypalmitate (16:0-OH)	3.22 (1.84-5.62)	<0.001	0.01	3.53 (2-6.22)	<0.001	0.00
5-Oxoproline	2.16 (1.38-3.40)	0.001	0.02	2.79 (1.73-4.51)	<0.001	0.00
Eicosanodioate (20:0-DC)	2.40 (1.29-4.45)	0.005	0.08	2.79 (1.5-5.21)	0.001	0.02
Eicosapentaenoate (EPA; 20:5)	1.76 (1.18-2.63)	0.006	0.08	2.02 (1.31-3.11)	0.001	0.02
Hydroxymyristate (14:0-OH)	2.69 (1.60-4.53)	<0.001	0.01	2.51 (1.39-4.54)	0.002	0.02
Hydroxystearate (18:0-OH)	2.65 (1.56-4.48)	<0.001	0.01	2.66 (1.55-4.56)	<0.001	0.01
Laurylcarnitine (C12:0)	2.63 (1.49-4.65)	0.001	0.02	2.62 (1.55-4.45)	<0.001	0.01
Linolenate (18:3)	2.81 (1.49-5.28)	0.001	0.03	3.59 (1.76-7.34)	<0.001	0.01
Linoleneoyl-carnitine (C18:3)	2.53 (1.32-4.86)	0.005	0.08	2.44 (1.22-4.88)	0.011	0.06
Myristate (14:0)	2.39 (1.28-4.47)	0.006	0.08	3.59 (1.8-7.16)	<0.001	0.01
Phenylalanyltryptophan	0.48 (0.28-0.81)	0.006	0.08	0.55 (0.32-0.97)	0.037	0.14

Cox regression models in future pancreatic cancer cases (n=82 for LCMS, n=81 for GCMS) with adjustment for lag-time to diagnosis and lag-time*metabolite ('Crude') and further adjustment for BMI, type of surgery, smoking status, fasting status, age, sex, storage time and TNM stage ('Adjusted'). HR, hazard ration per standard deviation increase; GCMS, gas chromatography mass spectrometry; LCMS, liquid chromatography mass spectrometry; FDR, false discovery rate.

Table S12 Conditional logistic regression models of branched chain amino acids (BCAA)

Metabolite	All		Males		Females	
	OR (95 % CI)	P value	OR (95 % CI)	P value	OR (95 % CI)	P value
Isoleucine	0.93 (0.65-1.32)	0.7	1.13 (0.6-2.14)	0.7	0.87 (0.58-1.31)	0.5
Leucine	0.82 (0.56-1.2)	0.3	0.95 (0.44-2.04)	0.9	0.80 (0.53-1.21)	0.3
Valine	1.18 (0.85-1.64)	0.3	1.25 (0.74-2.1)	0.4	1.13 (0.74-1.73)	0.6
BCAA[1]	0.88 (0.61-1.27)	0.5	1.11 (0.53-2.33)	0.8	0.84 (0.56-1.25)	0.4

[1], sum of isoleucine, leucine, and valine.

https://dx.doi.org/10.21037/jgo-23-930

References

46. Jiye A, Trygg J, Gullberg J, et al. Extraction and GC/MS analysis of the human blood plasma metabolome. Anal Chem 2005;77:8086-94.
47. Jonsson P, Johansson AI, Gullberg J, et al. High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses. Anal Chem 2005;77:5635-42.
48. Schauer N, Steinhauser D, Strelkov S, et al. GC-MS libraries for the rapid identification of metabolites in complex biological samples. FEBS Lett 2005;579:1332-7.

https://dx.doi.org/10.21037/jgo-23-930