BMC Medicine

RESEARCH

Open Access

Association of circulating metabolic biomarkers with risk of lung cancer: a population-based prospective cohort study

Lan Wu[1], Jun Yang[2], Yu Chen[1], Jiahao Lin[1], Wenkai Huang[3] and Mengmeng Li[1*]

Abstract

Background There is emerging evidence that metabolites might be associated with risk of lung cancer, but their relationships have not been fully characterized. We aimed to investigate the association between circulating metabolic biomarkers and lung cancer risk and the potential underlying pathways.

Methods Nuclear magnetic resonance metabolomic profiling was conducted on baseline plasma samples from 91,472 UK Biobank participants without cancer and pregnancy. Multivariate Cox regression models were employed to assess the hazard ratios (HRs) of 164 metabolic biomarkers (including metabolites and lipoprotein subfractions) and 9 metabolic biomarker principal components (PCs) for lung cancer, after adjusting for covariates and false discovery rate (FDR). Pathway analysis was conducted to investigate the potential metabolic pathways.

Results During a median follow-up of 11.0 years, 702 participants developed lung cancer. A total of 109 metabolic biomarkers (30 metabolites and 79 lipoprotein subfractions) were associated with the risk of lung cancer. Glycoprotein acetyls demonstrated a positive association with lung cancer risk [HR = 1.13 (95%CI: 1.04, 1.22)]. Negative associations with lung cancer were found for albumin [0.78 (95%CI: 0.72, 0.83)], acetate [0.91 (95%CI: 0.85, 0.97)], valine [0.90 (95%CI: 0.83, 0.98)], alanine [0.88 (95%CI: 0.82, 0.95)], glucose [0.91 (95%CI: 0.85, 0.99)], citrate [0.91 (95%CI: 0.85, 0.99)], omega-3 fatty acids [0.83 (95%CI: 0.77, 0.90)], linoleic acid [0.83 (95%CI: 0.77, 0.89)], etc. Nine PCs represented over 90% of the total variances, and among those with statistically significant estimates, PC1 [0.85 (95%CI: 0.80, 0.92)], PC2 [0.88 (95%CI: 0.82, 0.95)], and PC9 [0.87 (95%CI: 0.80, 0.93)] were negatively associated with lung cancer risk, whereas PC7 [1.08 (95%CI: 1.00, 1.16)] and PC8 [1.16 (95%CI: 1.08, 1.26)] showed positive associations with lung cancer risk. The pathway analysis showed that the "linoleic acid metabolism" was statistically significant after the FDR adjustment (*p* value 0.0496).

Conclusions Glycoprotein acetyls had a positive association with lung cancer risk while other metabolites and lipoprotein subfractions showed negative associations. Certain metabolites and lipoprotein subfractions might be independent risk factors for lung cancer. Our findings shed new light on the etiology of lung cancer and might aid the selection of high-risk individuals for lung cancer screening.

Keywords Metabolic biomarkers, Metabolites, Lipoprotein subfractions, Lung cancer

*Correspondence:
Mengmeng Li
limm@sysucc.org.cn
Full list of author information is available at the end of the article

Background

Lung cancer is one of the most common and fatal cancers in the world, with 2.5 million new cases of lung cancer and 1.8 million new lung cancer deaths in 2022 according to GLOBOCAN [1]. Lung cancer screening is the most effective strategy to address the rising global public health challenge [2]. Current lung cancer screening has considered the traditional risk factors to identify high-risk populations [3] and around one third of lung cancer cases would still be missed when these risk factors were used in the selection criteria [4]. Therefore, it is crucial to expand our knowledge regarding risk factors of lung cancer and to identify biomarkers that could potentially be used for lung cancer screening.

Numerous biomarkers can be quantified in a single measurement using metabolomic profiling, allowing subsequent quantification of the associations between metabolites and physiological and pathological changes, which captures the effects of genetic variation, environmental factors, and their interactions [5]. It has the potential to further enhance our understanding of the potential metabolic mechanisms of diseases [6, 7]. Metabolic biomarkers have become a pivotal part of scientific research, with applications ranging from risk prediction, causal analysis, gene discovery, and drug target validation [8–11]. Prospective studies have investigated associations between some circulating metabolic biomarkers (e.g., amino acids, fatty acids, glucose, lipids) and other diseases (diabetes and cardiovascular diseases) and found that metabolic biomarkers are closely related to disease occurrence [12, 13]. Few studies have explored the association between limited lipids and metabolites and the risk of lung cancer and obtained inconsistent results [14–16]. A few studies explored the association of metabolic biomarkers with multiple cancer types [17–20], but a detailed and systematic analysis on lung cancer, with lung cancer-specific covariates adequately adjusted, is still lacking, and potential metabolic mechanisms have not yet been explored.

Using recently available data from the UK Biobank (UKB) cohort, we characterize the associations of circulating metabolic biomarkers, quantified by a high-throughput targeted nuclear magnetic resonance (NMR) metabolomics platform, with the risk of incident lung cancer, and then explored the potential pathway of circulating metabolic biomarkers. This study can help us to have a deeper understanding on potential metabolic mechanisms of lung cancer and identify high-risk individuals that might benefit from lung cancer screening.

Methods

Study design and study population

The UKB is a multi-center, population-based, large-scale prospective cohort with 502,411 participants recruited [21]. The study design and study population details of the UKB have been described previously [22] and are available online (https://www.ukbiobank.ac.uk/). Briefly, adults aged 37–73 years were registered with the UK National Health Service between 2006 and 2010 from 22 assessment centers across England, Wales, and Scotland. Participants underwent an extensive range of baseline assessments that included self-administered touchscreen questionnaires, physical measurements, and biological samples. Then, they consented to track their health over time through linkage to electronic health records. The cancer registry data was available until 31 January 2021 for Scotland and 29 February 2020 for England and Wales. The UKB study received ethical approval from the National Health Service National Research Ethics Service (11/NW/0382; 16/NW/0274). All participants gave written informed consent.

In the present study, we excluded those (i) with prevalent cancers at baseline (except for non-melanoma skin cancer, coded as C44 using the International Classification of Diseases, 10th Revision [ICD-10]) ($N=37,940$), (ii) pregnant at baseline assessment ($N=147$), (iii) NMR metabolites data unavailable ($N=361,724$), (iv) with extreme metabolic biomarkers data, defined as outlying metabolic biomarker values outside the top or bottom 0.1% of the metabolic biomarker distribution ($N=8383$), and (v) with missing or outlying covariates data ($N=2745$) (Additional file 1: Fig. S1).

Plasma biomarker profiling by NMR

Metabolomic profiling was performed on baseline plasma samples from a randomly chosen subset of approximately 120,000 UKB participants [10] using a high-throughput NMR metabolomics platform [23, 24]. The blood sample handling and storage protocol has been previously described [25]. The quality control of metabolic biomarkers is available elsewhere [17]. Briefly, two internal control samples from Nightingale Health and two blind duplicate samples from UK Biobank were utilized to monitor consistency metrics throughout the project. Samples were measured using six NMR spectrometers. The coefficients of variation across the biomarker measures were below 5% for most biomarkers. Other metrics, including technical consistency of measurements over consecutive shipment batches and in different NMR spectrometers, the correlation of blinded duplicate samples for each biomarker, and the biological consistency

in repeat-visit samples drawn from the same individuals 4 years apart, were also assessed to ensure the technical and biological repeatability of the measurements. Additionally, comparisons of the NMR biomarker measurements to routine clinical chemistry are illustrated in Additional file 1: Fig. S2, with Pearson correlation coefficients > 0.8 for all biomarkers except for albumin.

This simultaneously quantified 249 metabolic biomarkers (164 directly measured with absolute concentrations and 4 with the unit of diameters or degree, and 81 ratios of these), including metabolites and lipoprotein subfractions. A subset of 164 metabolic biomarkers with directly measured concentrations was selected for inclusion in the presented analyses to enable comparisons across metabolic biomarkers in their effect estimates.

Covariates

Questionnaires collected information on sociodemographic and lifestyle factors, including age, sex, Townsend deprivation index (classified into five quartiles), and smoking status (never, former, and current smokers). Family history of lung cancer was collected (yes or no). Under standard protocols, physical measurements included height, weight, and spirometry, etc. Chronic obstructive pulmonary disease (COPD) was defined as a participant who met the spirometry criteria of having an FEV1/FVC ratio less than 0.7 (GOLD-COPD) and a percentage of the predicted value of FEV1 less than 80% or who had been diagnosed by a physician as having COPD (emphysema or chronic bronchitis) [26]. Body mass index (BMI) was calculated as weight/height^2 and further classified as underweight (less than 18.5 kg/m^2), normal (18.5–24.9 kg/m^2), overweight (25.0–29.9 kg/m^2), and obesity ($\geq$ 30 kg/m^2). BMI less than 12 or larger than 60 kg/m^2 were considered outliers and would be deleted. Fasting hours before taking biological samples were recorded and those less than 1 h or more than 6 h were grouped due to small number of participants.

Outcomes

The incident lung cancer cases were defined by ICD-10 codes (C33–C34). Due to the difficulties in distinguishing between metastatic and second primary cancer, as well as the varying risk profiles of lung cancer among individuals with a prior cancer diagnosis, participants were censored if cancer of any other site (apart from non-melanoma skin cancer) was identified before lung cancer. Every participant was tracked prospectively until one of the following events happened: diagnosis of lung cancer, censored, death, or lost follow-up. Primary lung cancers are classified as small-cell lung cancers (SCLC) and non-small cell lung cancers (NSCLC), with the latter consisting mostly of adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC) [27].

Statistical analysis

Baseline characteristics were presented using mean (and standard deviation, SD) for continuous variables and frequency (and percentage) for categorical variables. *T*-test, Wilcoxon rank-sum test, and χ^2 tests were used where appropriate to compare baseline characteristics between the NMR subset and total UKB participants. All NMR biomarkers were log-transformed and standardized.

The proportional hazard assumption was checked by tests based on Schoenfeld residuals. The Cox proportional hazards model yielded hazard ratios for the associations of 164 individual metabolic biomarkers after adjusting for age (continuous), sex (categorical, male versus female), Townsend deprive index (categorical, with the first quartile as referent), COPD (categorical, yes versus no), family history of lung cancer (categorical, yes versus no), smoking status (categorical, with never smokers as referent), BMI (categorical, with normal as referent), and fasting time (categorical, with $\leq$ 1 h as referent). To investigate the shape of the relationships, subjects were divided into baseline groups based on distribution quartiles, and a trend test was performed across quartiles (Additional file 1: Fig. S3). Continuous assessments of each NMR biomarker were conducted to determine the HR per 1 SD increase in the metabolic biomarkers. Pearson correlation coefficients were calculated to quantify the correlation between metabolic biomarkers and were presented in a heatmap (Additional file 1: Fig. S4).

Principal component analysis (PCA) is to project high-dimensional data into a low-dimensional space while retaining as much variance as possible [28, 29]. With the overall measure of sampling adequacy (MSA) of KMO test being 0.95 (> 0.80) and p value of Bartlett test < 0.001, the orthogonal PCA was further performed. And the number of principal components (PCs) is selected based on the scree plot with eigenvalue > 1 and cumulative proportion of variance $\geq$ 90%. Based on the corresponding result (Additional file 1: Fig. S5), a total of 9 PCs was selected and incorporated into the Cox hazard model to assess the associations between these PCs and risk of lung cancer. Finally, we calculated the loadings from PCs to understand the correlations between original metabolic biomarkers and PCs.

Selected individual metabolic biomarkers derived from the Cox regression model must meet the criteria of false discovery rate (FDR) adjusted p value < 0.05 and be mapped in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (those that cannot be mapped directly would be removed). We performed pathway analysis

by overrepresentation and pathway topology through MetaboAnalyst 6.0 [30]. The metabolic biomarkers with FDR adjusted p value < 0.05 were subject to mapping the metabolites in KEGG, then pathway analysis was conducted for the mapped metabolites in MetaboAnalyst. Metabolic pathways with impact > 0.2 were considered in the study [31].

Subgroup analyses were performed to explore the potential modifying effects of covariates on associations between metabolic biomarkers and lung cancer. Specifically, the primary analysis was extended to consider subgroups by histological subtype (NSCLC, SCLC; and for NSCLC, we also considered LUAD and LUSC), sex (female, male), and smoking status (never smoker, ever smoker).

We conducted several sensitivity analyses to determine the robustness of our findings. Firstly, we removed participants with a follow-up of less than 2 years to avoid reverse causality. Secondly, we removed those who fasted for less than 3 h to minimize the impact of fasting on metabolic indicators in plasma. Thirdly, we added pack years into the multivariate Cox models to minimize the residual confounding of smoking. Furthermore, we excluded participants diagnosed with cancer (except for non-melanoma skin cancer) shortly after baseline (i.e., ≤ 3 or ≤ 5 years) to reduce the influence of delayed cancer diagnosis on the results.

All statistical analyses were performed with the R software (version 4.3.1, R foundation for statistical computing, Vienna, Austria). Statistical significance was defined as a 2-sided p value < 0.05.

Results

Of the total 502,411 UKB participants, a random set of approximately 118,000 (23%) participants had the NMR-measured metabolic biomarkers data. After meeting the inclusion criteria (Additional file 1: Fig. S1), 91,472 participants were included in this study. The mean age of participants in this study was 56.3 years (standard deviation: 8.1 years) and 53% were women (Table 1). The study participants' baseline characteristics are similar to those of the total UKB participants (Additional file 1: Table S1).

After a total of 964,292 person-years of follow-up (median 11.0 years, mean 10.5 years), 702 participants developed primary lung cancer. The incidence rate of lung cancer was 72.8 per 100,000. In multivariate Cox models, 109 of 164 metabolic biomarkers had statistically significant relationships with the risk of lung cancer (FDR adjusted p value < 0.05) (Fig. 1, Additional file 1: Table S2). Glycoprotein acetyls, a marker of inflammation, demonstrated a positive association with lung cancer risk, with an HR of 1.13 (95%CI: 1.04, 1.22). Albumin [0.78 (95%CI: 0.72, 0.83)], acetate [0.91 (95%CI: 0.85, 0.97)], valine [0.90 (95%CI: 0.83, 0.98)], histidine [0.90 (95%CI: 0.84, 0.97)], alanine [0.88 (95%CI: 0.82, 0.95)], glucose [0.91 (95%CI: 0.85, 0.99)], citrate [0.91 (95%CI: 0.85, 0.99)], omega-3 fatty acids [0.83 (95%CI: 0.77, 0.90)], apolipoprotein A1 [0.86 (95%CI: 0.78, 0.93)], apolipoprotein B [0.86 (95%CI: 0.82, 0.94)], fatty acids (HRs around 0.81–0.90), glyceride phospholipids (HRs around 0.82–0.86), and lipids (HRs around 0.83–0.86) were negatively associated with lung cancer risk (Fig. 1A). In addition, all lipoprotein subfractions are related to negative lung cancer risk, with the HRs ranging from 0.83 to 0.92 (Fig. 1B).

The top 9 PCs of the metabolic biomarkers accounted for 90.7% of the total variance in the 164 individual biomarkers (Additional file 1: Fig. S5). As shown in Fig. 2, the principal components' loadings could be observed (the higher a biomarker's loading, the more it contributes to that PC). The major contributors in the first 4 PCs

Table 1 Baseline characteristics of the study population

Characteristic	Study population (n = 91,472)
Age, years, mean (SD)	56.3 (8.1)
Sex (%)	
Women	48,385 (52.9%)
Men	43,087 (47.1%)
Townsend deprivation index, mean (SD)	− 1.4 (3.1)
Smoking status (%)	
Never	50,601 (55.3%)
Former	31,494 (34.4%)
Current	9377 (10.2%)
Family history of lung cancer (%)	
No	79,999 (87.5%)
Yes	11,473 (12.5%)
COPD (%)	
No	86,016 (94.0%)
Yes	5456 (6.0%)
BMI category (%)	
Underweight	419 (0.5%)
Normal	29,365 (32.1%)
Overweight	39,390 (43.1%)
Obesity	22,298 (24.4%)
Fasting time, hour (%)	
≤ 1	4417 (4.8%)
2	19,252 (21.1%)
3	26,800 (29.3%)
4	20,075 (21.9%)
5	10,977 (12.0%)
≥ 6	9951 (10.9%)

SD standard deviation, *COPD* chronic obstructive pulmonary disease, *BMI* body mass index

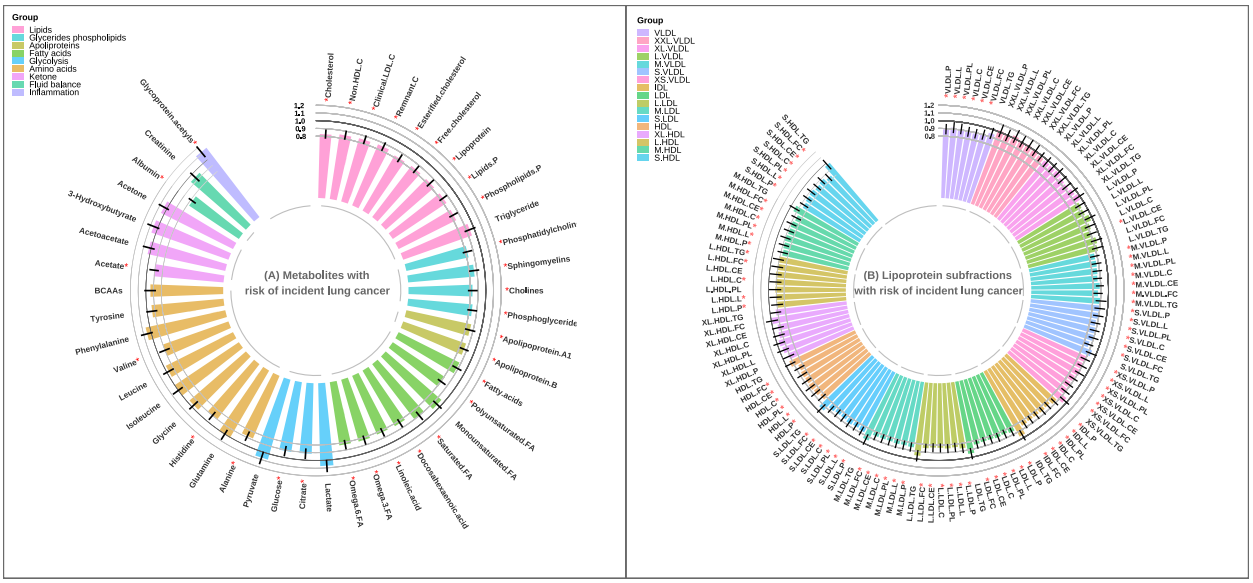

Fig. 1 Associations of metabolic biomarkers with risk of lung cancer among the study population. Hazard ratios (with 95% confidence intervals) are presented per 1 – SD higher metabolic biomarker on the natural log scale, adjusted for age, sex, Townsend deprivation index, smoking, COPD, family history of lung cancer, fasting hours, and body mass index. *FDR adjustment *p* value < 0.05. FA, fatty acids; BCAAs, branched-chain amino acids; HDL, high-density lipoproteins; IDL, intermediate-density lipoproteins; LDL, low-density lipoproteins; VLDL, very low-density lipoproteins; XXL, extremely large; XL, very large; L, large; M, medium; S, small; XS, very small; P, lipoprotein particle; L, total lipid; PL, phospholipid; C, cholesterol; CE, esterified cholesterol; FC, free cholesterol; TG, triglyceride

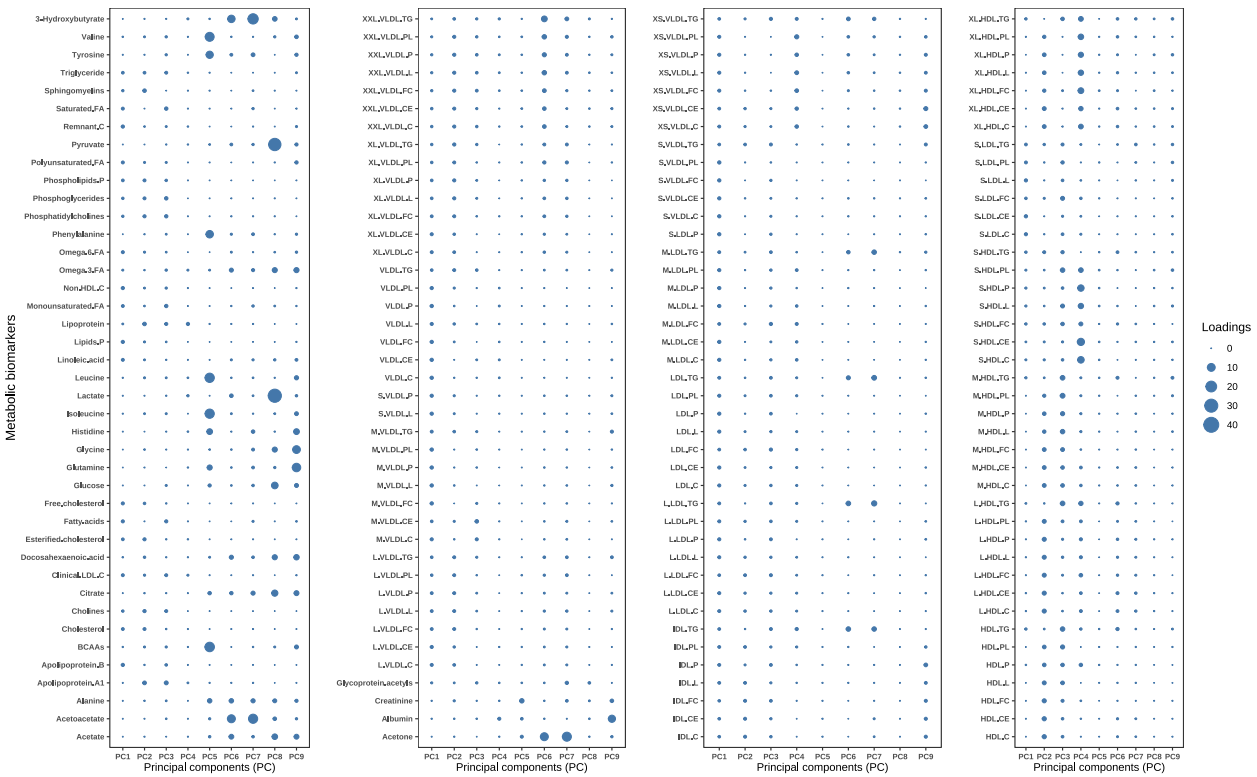

Fig. 2 The factor loadings of principal components

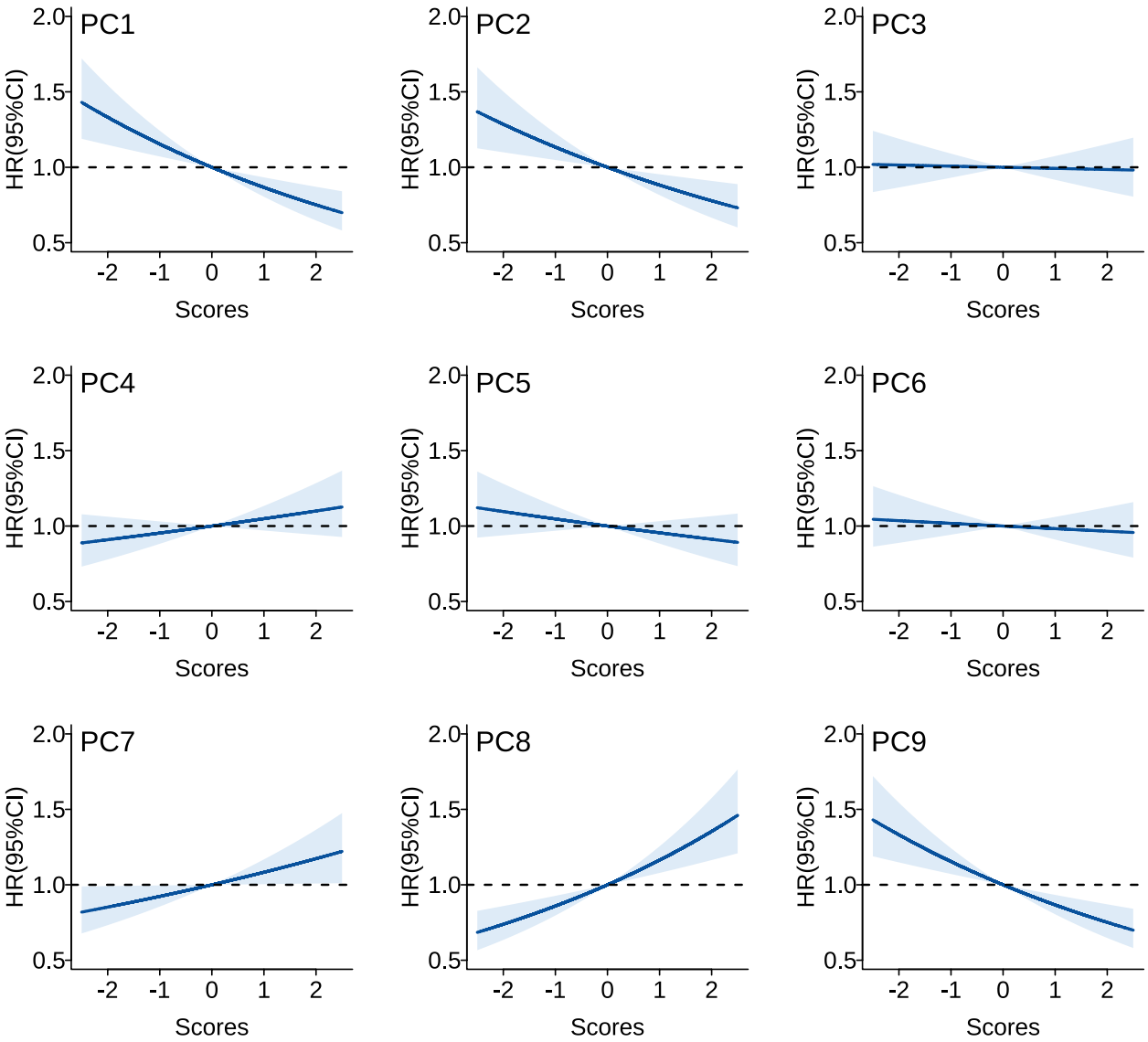

Fig. 3 Associations of principal components with risk of lung cancer among the study population. Adjusted for age, sex, Townsend deprivation index, smoking, COPD, family history of lung cancer, fasting hours, and body mass index

were lipoprotein subfraction concentrations, while for the last 5 PCs, the metabolite concentrations were prominent contributors.

The relationships of PCs with lung cancer are presented in Additional file 1: Table S3 and Fig. 3. A statistically significant association was found between 5 out of 9 PCs and risk of lung cancer. Pyruvate, lactate, acetone, citrate, and glucose predominate in PC8 [1.16 (95%CI: 1.08, 1.26)] and glycoprotein acetyls, 3-hydroxybutyrate, acetoacetate, and acetone predominate in PC7 [1.08 (95%CI: 1.00, 1.16)], both exhibited a positive association with lung cancer. Albumin, glutamine, glycine, and histidine contributed to PC9 [0.87 (95%CI: 0.80, 0.93)] and cholesterol, lipids, fatty acids, apolipoproteins, and lipoprotein subfractions prominently contributed to both PC1 [0.87 (95%CI: 0.81, 0.93)] and PC2 [0.88 (95%CI: 0.82, 0.95)], showing negative associations with lung cancer.

Results of the pathway analysis are shown in Fig. 4 and Additional file 1: Table S4. A total of 25 pathways were related to the 14 metabolites which were significantly associated with lung cancer risk. There are three pathways with an impact > 0.2. The first pathway was the "linoleic acid metabolism" (FDR adjusted p value: 0.0496), with 5 total compounds including 2 Hits corresponding to linoleic acid and phosphatidylcholine. The second

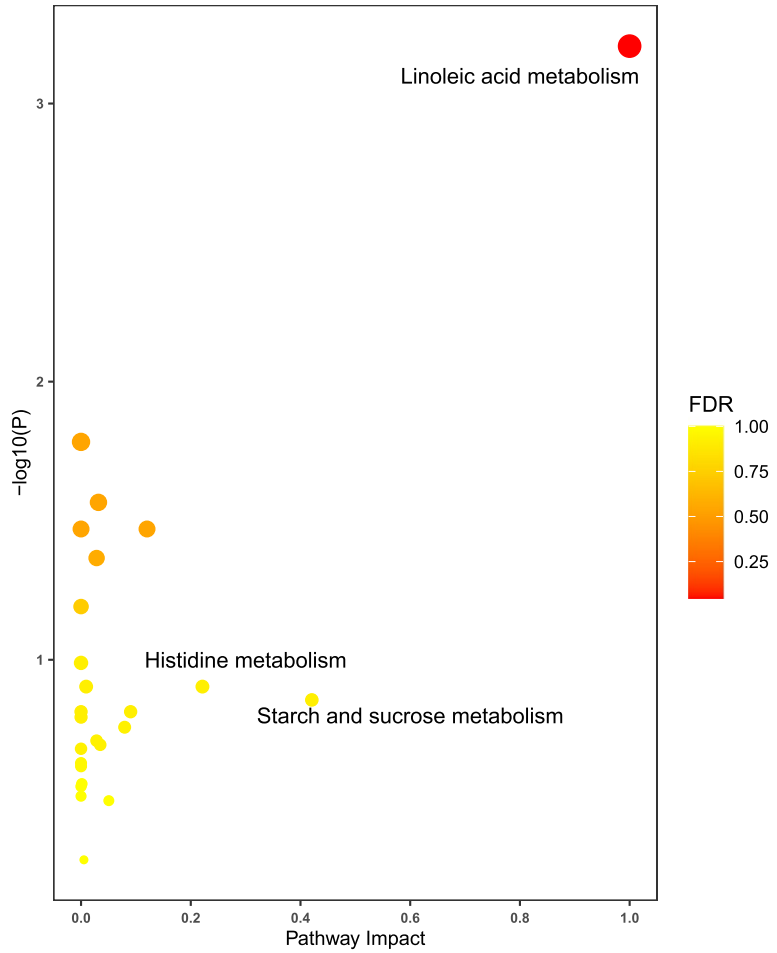

Fig. 4 Summary of pathways analysis using MetaboAnalyst

pathway was the "starch and sucrose metabolism" (FDR adjusted p value: 0.92). This pathway included 18 total compounds among which there was 1 Hit corresponding to glucose. The third pathway was the "histidine metabolism" (FDR adjusted p value: 0.92), which included 16 total compounds among which there was 1 Hit corresponding to histidine.

Of the 702 primary lung cancer cases, 425 were classified as NSCLC (including 275 LUAD and 131 LUSC cases) and 82 as SCLC. Eighty-nine metabolic biomarkers were associated with NSCLC and 30 metabolic biomarkers were associated with LUAD after FDR adjustment. Glycoprotein acetyls remained the only significant metabolic biomarker that is positively associated with the risk of lung cancer in the NSCLC group, while albumin, choline, and lipoprotein subfractions were negatively associated with the risk of lung cancer in both NSCLC and LUAD groups. And albumin and omega-3 fatty acids were inversely associated with LUSC after FDR adjustment. We did not find any metabolic biomarker that was associated with SCLC (Additional file 1: Fig. S6). Among females, we observed that glycoprotein acetyls and lactate increased the risk of lung cancer, but cholesterol, alanine, and lipoprotein subfraction were adversely associated with lung cancer. A total of 109 metabolic biomarkers were found to be negatively associated with lung cancer risk in the male population. Among smokers, 96 metabolic biomarkers were associated with lung cancer risk after FDR adjustment, with glycoprotein acetyls, lactate, and pyruvate showing positive associations and other metabolic biomarkers presenting negative associations. Null findings were observed for non-smokers (Additional file 1: Fig. S7).

In the sensitivity analysis, results were broadly similar to the main findings. The positive association between glycoprotein acetyls and lung cancer remained after adjusting for fasting time; and there were 79 metabolic biomarkers negatively associated with lung cancer risk after adjusting for follow-up time, fasting time, and pack years (Fig. 5). In addition, 93 metabolic biomarkers

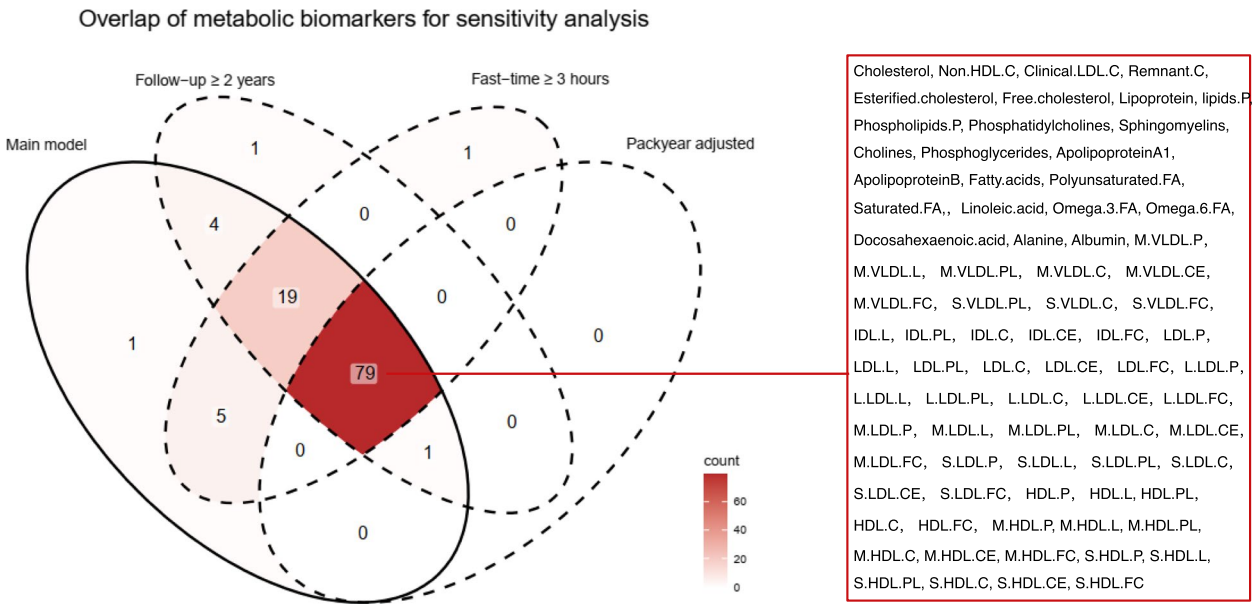

Fig. 5 Overlap of metabolic biomarkers for sensitivity analysis. FA, fatty acids; HDL, high-density lipoproteins; IDL, intermediate density lipoproteins; LDL, low-density lipoproteins; VLDL, very low-density lipoproteins; XXL, extremely large; XL, very large; L, large; M, medium; S, small; XS, very small; P, lipoprotein particle; L, total lipid; PL, phospholipid; C, cholesterol; CE, esterified cholesterol; FC, free cholesterol; TG, triglyceride

remained statistically significant in the associations with lung cancer risk after excluding participants diagnosed with cancer shortly after baseline (≤ 3 and ≤ 5 years) (Additional file 1: Fig. S8).

Discussion

This prospective population-based cohort study among 91,472 participants explored the association of NMR-measured metabolic biomarkers (metabolites, lipoprotein subfractions) with lung cancer risk; 109 of 164 metabolic biomarkers had statistically significant relationships with lung cancer. Glycoprotein acetyls demonstrated a robust and positive association with lung cancer risk, with the highest hazard ratio compared to other metabolites. Negative associations with lung cancer were found for albumin, acetate, valine, histidine, alanine, glucose, citrate, fatty acids, apolipoproteins, glyceride phospholipids, lipids, and lipoprotein subfractions (around 0.8–0.9). The pathway analysis showed that the "linoleic acid metabolism" was statistically significant after applying the FDR adjustment (p value 0.0496). To our knowledge, this study is the first to systematically evaluate the impact of metabolic biomarkers on the risk of lung cancer based on a large-scale cohort. Our study can deepen the understanding on potential metabolic mechanisms of lung cancer and help to identify high-risk individuals that might benefit from lung cancer screening.

Inflammation predisposes the development of cancer and accelerates all stages of tumorigenesis [32, 33], with glycoprotein acetyls being a representative biomarker. Our study found a strong positive association of glycoprotein acetyls with lung cancer risk, which is consistent with prior studies on glycoprotein acetyls and lung cancer risk. A strong association was observed between glycoprotein acetyls and cardiovascular disease, T2D, COPD, and lung cancer in population-based cohort studies in Finland and the UK [34, 35]. Our study also observed a positive association of lactate and pyruvate with lung cancer risk, though marginally significant. Metabolic intermediates of central carbon metabolism were found to be signaling molecules that control hypoxic signals and stress response [19]. Lactate binds to and stabilizes NDRG family member 3 (NDRG3) to enhance Raf-ERK1/2 signaling and promotion of angiogenesis [36]. Moreover, lactate and pyruvate were glycolytic by-products and oncometabolites, which can modulate immune cell function, creating an immunosuppressive microenvironment that favors tumor progression and initiates or sustains tumor growth and metastasis [37–39]. According to prior studies, lactate contributed to central metabolism in human NSCLC in vivo [40, 41], and pyruvate was rapidly taken up by malignant tissue [42]; glycolysis-related metabolites are closely related to the occurrence and development of lung cancer.

Albumin was associated with a lower lung cancer risk in our study. Albumin, an important component in fluid balance, has proven to be closely associated with the progression and prognosis of cancer [43, 44]. As an

endogenous antioxidant, albumin may reduce cancer risk through exerting anticarcinogenic properties [45]. Some epidemiological studies found that albumin is positively associated with lung function, suggesting that people with low albumin level may have poor lung function [46], which is opposite to glycoprotein acetyls [35]. The association between albumin and lung cancer risk may also reflect systemic inflammation [47]. Amino acids are regulators of cancer stem cells. Low valine to isoleucine/leucine ratio, or BCAA imbalance, slows the formation of hematopoietic stem cells [48]. Valine also showed a negative association with lung cancer risk in our study, but a positive association was found with squamous cell lung cancer in another study [49]. The association between valine and lung cancer risk was inconsistent and should be further studied.

Lipids, which serve as energy resources and components of the cell membrane, have immunomodulatory functions and the potential to influence cancer immunity [50]. Our study found most lipoprotein subfractions were negatively associated with lung cancer risk, especially HDL particles. Both NMR-measured and clinical chemistry measured HDL-C were inversely associated with the risk of lung cancer [17], which is consistent with previous studies [14, 51]. HDL could contribute to lung cancer carcinogenesis through its role in cholesterol reverse transportation, regulating inflammatory and proliferative pathways [52]. Meanwhile, recent studies showed that both low HDL-C and low apolipoprotein A are associated with an increased lung cancer risk [51, 53]. The consistent findings on fatty acids, lipids, and lipoproteins imply the potential protective role of these factors.

In our study, fatty acids were observed to be associated with a low lung cancer risk. Another study also found that polyunsaturated fatty acids exhibited an inverse association with lung cancer and showed an additive interaction with genetic risk [16], supporting that polyunsaturated fatty acids may serve as protective factors for lung cancer. Omega-3 fatty acids can modulate epigenetic events to regulate cellular processes associated with carcinogenesis [54]. It has been found that omega-3 fatty acids directly upregulates 15-prostaglandin dehydrogenase (15-PGDH), which acts as a tumor suppressor in lung and colon cancer [55, 56], which is consistent with our results. In addition, the electrophilic oxidized derivatives of omega-3 fatty acids regulated catalytic histone modification or DNA methylation of enzymes to control miRNA expression [57], and docosahexaenoic acid inhibits angiogenesis by triggering exosome secretion of miRNAs which promotes the expression of angiogenic genes in endothelial cells [58], which provides the evidence that omega-3 fatty acids and docosahexaenoic acid can lower the risk of cancer [59]. Omega-6 fatty acids and linoleic acid may possess anti-cancer effects by disrupting the cell cycle in the G1 phase, upregulating the protein expression of the cell cycle inhibitor p21, and decreasing the expression of cyclins A and D [60]. Recent studies have demonstrated that linoleic acid has the capacity to enhance CD8 + T cell metabolism, prevent exhaustion, and stimulate memory-like phenotypes with superior effector functions, resulting in greater antitumor potency in vitro and in mouse model [61]. Besides, linoleic acid as a biomarker to predict lung cancer has a good performance in squamous cell carcinoma and adenocarcinoma [62].

Based on the pathway analysis, "linoleic acid metabolism" emerged as the most potential metabolic process and remained statistically significant after FDR adjustment. Recent studies demonstrated that lipid metabolism plays an important role in cancer biology by influencing cell growth, survival, proliferation, migration, invasion, and metastasis [63]. And metabolic alterations in "linoleic acid metabolism" metabolic pathways can be observed in NSCLC [64, 65]. Changes in lipids in the primary tumor microenvironment play a critical role in facilitating carcinogenesis, escape, and spread, as well as evading immune surveillance. Significant enhancements in "linoleic acid metabolism" were observed in the lung microbiome in mice with lung cancer [66]. Our pathway analyses also showed a possible role of the "starch and sucrose metabolism." Starch and sucrose metabolism genes were shown to be highly upregulated in metastatic cancer cell lines [67]. In a paper from Yang et al. [68], the occurrence of lung squamous cell carcinoma may be regulated by the genes AMY2B and AMY2A via the "starch and sucrose metabolism" pathway. These metabolic pathways need to be further verified in experimental studies.

PCA is frequently used in metabolomics analysis because it simplifies the complexity of high-dimensional data while retaining trends and patterns [69]. The PCs aggregated signatures of metabolic biomarkers across multiple molecular pathways, representing a large number of highly correlated metabolic biomarkers. The association of each principal component with lung cancer risk is consistent with its major contributor's metabolic biomarkers. The association of individual metabolic biomarkers with the risk of lung cancer has been maintained after reduced dimension, which again demonstrates the robustness of our results.

Our study has several strengths. Using an established targeted NMR metabolomics platform, with existing clinical regulatory approvals, enabled the quantification of diverse biomarkers and enhanced the potential clinical relevance [10]. Moreover, high levels of correlation between NMR- and standard clinical chemistry-derived concentrations of a subset of biomarkers (Additional

file 1: Fig. S2) support the validity of the approach [70]. Importantly, these associations were robust after controlling for a wide range of potential confounders and in a series of sensitivity analyses.

Several limitations of this study should be noted. First, we used only one single blood sample and did not consider the changes of metabolic biomarkers over time, which may overlook the dynamics of metabolic biomarkers. Second, blood samples were taken in the non-fasting state, and therefore would be subject to greater variability in metabolic biomarker concentrations than fasting samples. However, our analyses were adjusted for fasting time, which should have limited any material impact of the use of non-fasting samples on the findings. It is also reassuring that in the sensitivity analysis, the results were robust when we excluded those with short fasting times. Third, in the stratified analyses, null findings were observed for non-smokers and those with SCLC, which might be partly due to the insufficiency in statistical power in these strata. Fourth, pathway analysis cannot be conducted for some metabolic biomarkers because they could not be mapped in KEGG, which may lead to loss of information on some pathways that may impact the lung cancer risk.

Conclusions

In this study, 109 metabolic biomarkers were associated with the risk of lung cancer. Glycoprotein acetyls had a positive association with lung cancer risk while other metabolites and lipoprotein subfractions showed negative associations. Certain metabolites and lipoprotein subfractions might be independent risk factors for lung cancer. Our findings shed new light on the etiology of lung cancer and might be useful in the selection of high-risk individuals for lung cancer screening.

Abbreviations

UKB	UK Biobank
NMR	Nuclear magnetic resonance
SD	Standard deviation
HR	Hazard ratio
95%CI	95% Confidence interval
BMI	Body mass index
COPD	Chronic obstructive pulmonary disease
NSCLC	Non-small cell lung cancers
SCLC	Small-cell lung cancers
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
FDR	False discovery rate
PCA	Principal component analysis
PC	Principal component
KEGG	Kyoto Encyclopedia of Genes and Genomes
FA	Fatty acids
BCAAs	Branched-chain amino acids
HDL	High-density lipoproteins
IDL	Intermediate-density lipoproteins
LDL	Low-density lipoproteins
VLDL	Very low-density lipoproteins
XXL	Extremely large
XL	Very large
L	Large
M	Medium
S	Small
XS	Very small
P	Lipoprotein particle
L	Total lipid
PL	Phospholipid
C	Cholesterol
CE	Esterified cholesterol
FC	Free cholesterol
TG	Triglyceride

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-025-03993-4.

Additional file 1: Table S1 Baseline characteristics of the study population and total UKB population. Table S2 Distribution of metabolic biomarkers and their associations with lung cancer among the study population. Table S3 Association of the first 9 metabolic biomarker principal components with risk of lung cancer among the study population. Table S4 Detailed results of pathways analysis. Fig. S1 Selection criteria of the study population. Fig. S2 Comparisons of the NMR biomarker measurements to routine clinical chemistry. Fig. S3 Associations of metabolic biomarkers with risk of lung cancer among the study population (by quartile). Fig. S4 Correlation between metabolic biomarkers. Fig. S5 Scree plot of metabolic biomarkers in principal component analysis. Fig. S6 Association of metabolic biomarkers with risk of lung cancer by histological subtype. Fig. S7 Associations of metabolic biomarkers with risk of lung cancer stratified by sex and smoking status. Fig. S8 Overlap of metabolic biomarkers in sensitivity analysis (excluding participants diagnosed with cancer shortly after baseline).

Acknowledgements

This study was conducted under the UK Biobank application number 91354. The authors gratefully acknowledge the participants and those involved in building the resource.

Authors' contributions

ML designed the study. LW performed the statistical analyses and drafted the manuscript. JY, YC, JL, and WH contributed to the interpretation of data and critically revised the manuscript. All authors have read and approved the final manuscript for publication.

Funding

This project was supported by Guangdong Basic and Applied Basic Research Foundation (No. 2025A1515011735), National Natural Science Foundation of China (No. 82204131), Guangzhou Science and Technology Project (No. 2025A04J3327), and Young Talents Program of Sun Yat-sen University Cancer Center (No. YTP-SYSUCC-0048). The funding sources were not involved in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Data availability

UK Biobank information is available online at the webpage (www.ukbiobank.co.uk). Data are available on application.

Declarations

Ethics approval and consent to participate

UK Biobank received ethical approval from the National Health Service National Research Ethics Service (11/NW/0382; 16/NW/0274) and was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent prior to any data collection.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Cancer Prevention, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-Sen University Cancer Center, Guangzhou, China. [2]School of Public Health, Guangzhou Medical University, Guangzhou, China. [3]National Central Cancer Registry Office, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.

Received: 15 August 2024 Accepted: 11 March 2025
Published online: 26 March 2025

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–63.
2. Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. Nat Rev Clin Oncol. 2023;20(9):624–39.
3. Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. Eur Respir J. 2016;48(3):889–902.
4. Robbins HA, Alcala K, Swerdlow AJ, Schoemaker MJ, Wareham N, Travis RC, et al. Comparative performance of lung cancer risk models to define lung screening eligibility in the United Kingdom. Br J Cancer. 2021;124(12):2026–34.
5. Yu L, Li K, Zhang X. Next-generation metabolomics in lung cancer diagnosis, treatment and precision medicine: mini review. Oncotarget. 2017;8(70):115774–86.
6. Pietzner M, Stewart ID, Raffler J, Khaw KT, Michelotti GA, Kastenmüller G, et al. Plasma metabolites to profile pathways in noncommunicable disease multimorbidity. Nat Med. 2021;27(3):471–9.
7. Karczewski KJ, Snyder MP. Integrative omics for health and disease. Nat Rev Genet. 2018;19(5):299–310.
8. Fang S, Holmes MV, Gaunt TR, Davey Smith G, Richardson TG. Constructing an atlas of associations between polygenic scores from across the human phenome and circulating metabolic biomarkers. Elife. 2022;11:e73951.
9. Borges MC, Haycock PC, Zheng J, Hemani G, Holmes MV, Davey Smith G, et al. Role of circulating polyunsaturated fatty acids on cardiovascular diseases risk: analysis using Mendelian randomization and fatty acid genetic association data from over 114,000 UK Biobank participants. BMC Med. 2022;20(1):210.
10. Julkunen H, Cichońska A, Slagboom PE, Würtz P. Metabolic biomarker profiling for identification of susceptibility to severe pneumonia and COVID-19 in the general population. Elife. 2021;10: e63033.
11. Shen M, Shi L, Xing M, Jiang H, Ma Y, Ma Y, et al. Unravelling the metabolic underpinnings of gestational diabetes mellitus: a comprehensive Mendelian randomisation analysis identifying causal metabolites and biological pathways. Diabetes Metab Res Rev. 2024;40(6): e3839.
12. Guasch-Ferré M, Hruby A, Toledo E, Clish CB, Martínez-González MA, Salas-Salvadó J, et al. Metabolomics in prediabetes and diabetes: a systematic review and meta-analysis. Diabetes Care. 2016;39(5):833–46.
13. Bragg F, Trichia E, Aguilar-Ramirez D, Bešević J, Lewington S, Emberson J. Predictive value of circulating NMR metabolic biomarkers for type 2 diabetes risk in the UK Biobank study. BMC Med. 2022;20(1):159.
14. Lin X, Lu L, Liu L, Wei S, He Y, Chang J, et al. Blood lipids profile and lung cancer risk in a meta-analysis of prospective cohort studies. J Clin Lipidol. 2017;11(4):1073–81.
15. Li M, Cao SM, Dimou N, Wu L, Li JB, Yang J. Association of metabolic syndrome with risk of lung cancer: a population-based prospective cohort study. Chest. 2024;165(1):213–23.
16. Huang J, Li J, He L, Miao J, Zhu M, Dai J, et al. The association between plasma fatty acids and risk of lung cancer: a prospective cohort study of the UK Biobank. J Clin Endocrinol Metab. 2024:dgae609. https://doi.org/10.1210/clinem/dgae609.
17. Julkunen H, Cichońska A, Tiainen M, Koskela H, Nybo K, Mäkelä V, et al. Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. Nat Commun. 2023;14(1):604.
18. Hu C, Fan Y, Lin Z, Xie X, Huang S, Hu Z. Metabolomic landscape of overall and common cancers in the UK Biobank: a prospective cohort study. Int J Cancer. 2024;155(1):27–39.
19. Wang YP, Li JT, Qu J, Yin M, Lei QY. Metabolite sensing and signaling in cancer. J Biol Chem. 2020;295(33):11938–46.
20. Papier K, Atkins JR, Tong TYN, Gaitskell K, Desai T, Ogamba CF, et al. Identifying proteomic risk factors for cancer using prospective and exome analyses of 1463 circulating proteins and risk of 19 cancers in the UK Biobank. Nat Commun. 2024;15(1):4010.
21. Conroy MC, Lacey B, Bešević J, Omiyale W, Feng Q, Effingham M, et al. UK Biobank: a globally important resource for cancer research. Br J Cancer. 2023;128(4):519–27.
22. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3): e1001779.
23. Würtz P, Kangas AJ, Soininen P, Lawlor DA, Davey Smith G, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in large-scale epidemiology: a primer on -omic technologies. Am J Epidemiol. 2017;186(9):1084–96.
24. Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. Circ Cardiovasc Genet. 2015;8(1):192–206.
25. Elliott P, Peakman TC. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. Int J Epidemiol. 2008;37(2):234–44.
26. Adeloye D, Song P, Zhu Y, Campbell H, Sheikh A, Rudan I. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. Lancet Respir Med. 2022;10(5):447–58.
27. Tang M, Abbas HA, Negrao MV, Ramineni M, Hu X, Hubert SM, et al. The histologic phenotype of lung cancers is associated with transcriptomic features rather than genomic characteristics. Nat Commun. 2021;12(1):7081.
28. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. Philos Trans A Math Phys Eng Sci. 2016;374(2065):20150202.
29. Ringnér M. What is principal component analysis? Nat Biotechnol. 2008;26(3):303–4.
30. Pang Z, Lu Y, Zhou G, Hui F, Xu L, Viau C, et al. MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. Nucleic Acids Res. 2024;52(W1):W398-w406.
31. Maan K, Baghel R, Dhariwal S, Sharma A, Bakhshi R, Rana P. Metabolomics and transcriptomics based multi-omics integration reveals radiation-induced altered pathway networking and underlying mechanism. NPJ Syst Biol Appl. 2023;9(1):42.
32. Greten FR, Grivennikov SI. Inflammation and cancer: triggers, mechanisms, and consequences. Immunity. 2019;51(1):27–41.
33. Karin M, Greten FR. NF-kappaB: linking inflammation and immunity to cancer development and progression. Nat Rev Immunol. 2005;5(10):749–59.
34. Kettunen J, Ritchie SC, Anufrieva O, Lyytikäinen LP, Hernesniemi J, Karhunen PJ, et al. Biomarker glycoprotein acetyls is associated with the risk of a wide spectrum of incident diseases and stratifies mortality risk in angiography patients. Circ Genom Precis Med. 2018;11(11): e002234.
35. Guo Y, Liu Q, Zheng Z, Qing M, Yao T, Wang B, et al. Genetic association of inflammatory marker GlycA with lung function and respiratory diseases. Nat Commun. 2024;15(1):3751.
36. Lee Dong C, Sohn Hyun A, Park ZY, Oh S, Kang Yun K, Lee KM, et al. A lactate-induced response to hypoxia. Cell. 2015;161(3):595–609.
37. Tauffenberger A, Fiumelli H, Almustafa S, Magistretti PJ. Lactate and pyruvate promote oxidative stress resistance through hormetic ROS signaling. Cell Death Dis. 2019;10(9):653.
38. Wishart DS. Is cancer a genetic disease or a metabolic disease? EBioMedicine. 2015;2(6):478–9.

39. Klein K, He K, Younes AI, Barsoumian HB, Chen D, Ozgen T, et al. Role of mitochondria in cancer immune evasion and potential therapeutic approaches. Front Immunol. 2020;11: 573326.
40. Faubert B, Li KY, Cai L, Hensley CT, Kim J, Zacharias LG, et al. Lactate metabolism in human lung tumors. Cell. 2017;171(2):358-71.e9.
41. Bonuccelli G, Tsirigos A, Whitaker-Menezes D, Pavlides S, Pestell RG, Chiavarina B, et al. Ketones and lactate "fuel" tumor growth and metastasis: evidence that epithelial cancer cells use oxidative mitochondrial metabolism. Cell Cycle. 2010;9(17):3506–14.
42. Nelson SJ, Kurhanewicz J, Vigneron DB, Larson PE, Harzstark AL, Ferrone M, et al. Metabolic imaging of patients with prostate cancer using hyperpolarized [1–13C]pyruvate. Sci Transl Med. 2013;5(198):198ra08.
43. Phillips A, Shaper AG, Whincup PH. Association between serum albumin and mortality from cardiovascular disease, cancer, and other causes. Lancet. 1989;2(8677):1434–6.
44. Xu SS, Li S, Xu HX, Li H, Wu CT, Wang WQ, et al. Haemoglobin, albumin, lymphocyte and platelet predicts postoperative survival in pancreatic cancer. World J Gastroenterol. 2020;26(8):828–38.
45. Ching S, Ingram D, Hahnel R, Beilby J, Rossi E. Serum levels of micronutrients, antioxidants and total antioxidant status predict risk of breast cancer in a case control study. J Nutr. 2002;132(2):303–6.
46. Ju S, Jeong JH, Heo M, Heo IR, Kim TH, Kim HC, et al. Serum albumin is a predictor of respiratory hospitalization in patients with bronchiectasis. Chron Respir Dis. 2021;18: 14799731211017548.
47. Yang Z, Zheng Y, Wu Z, Wen Y, Wang G, Chen S, et al. Association between pre-diagnostic serum albumin and cancer risk: results from a prospective population-based study. Cancer Med. 2021;10(12):4054–65.
48. Wilkinson AC, Morita M, Nakauchi H, Yamazaki S. Branched-chain amino acid depletion conditions bone marrow for hematopoietic stem cell transplantation avoiding amino acid imbalance-associated toxicity. Exp Hematol. 2018;63:12-6.e1.
49. Xu H, Wang X, Xu X, Liu L, Zhang Y, Yan X, et al. Association of plasma branched-chain amino acid with multiple cancers: a Mendelian randomization analysis. Clin Nutr. 2023;42(12):2493–502.
50. Karayama M, Inui N, Inoue Y, Yoshimura K, Mori K, Hozumi H, et al. Increased serum cholesterol and long-chain fatty acid levels are associated with the efficacy of nivolumab in patients with non-small cell lung cancer. Cancer Immunol Immunother. 2022;71(1):203–17.
51. Pedersen KM, Çolak Y, Bojesen SE, Nordestgaard BG. Low high-density lipoprotein and increased risk of several cancers: 2 population-based cohort studies including 116,728 individuals. J Hematol Oncol. 2020;13(1):129.
52. Pommier AJC, Alves G, Viennois E, Bernard S, Communal Y, Sion B, et al. Liver X receptor activation downregulates AKT survival signaling in lipid rafts and induces apoptosis of prostate cancer cells. Oncogene. 2010;29(18):2712–23.
53. von Eckardstein A, Nordestgaard BG, Remaley AT, Catapano AL. High-density lipoprotein revisited: biological functions and clinical relevance. Eur Heart J. 2023;44(16):1394–407.
54. Amarasekera M, Noakes P, Strickland D, Saffery R, Martino DJ, Prescott SL. Epigenome-wide analysis of neonatal CD4(+) T-cell DNA methylation sites potentially affected by maternal fish oil supplementation. Epigenetics. 2014;9(12):1570–6.
55. Huang G, Eisenberg R, Yan M, Monti S, Lawrence E, Fu P, et al. 15-Hydroxyprostaglandin dehydrogenase is a target of hepatocyte nuclear factor 3beta and a tumor suppressor in lung cancer. Cancer Res. 2008;68(13):5040–8.
56. Myung SJ, Rerko RM, Yan M, Platzer P, Guda K, Dotson A, et al. 15-Hydroxyprostaglandin dehydrogenase is an in vivo suppressor of colon tumorigenesis. Proc Natl Acad Sci U S A. 2006;103(32):12098–102.
57. Ravindra KC, Narayan V, Lushington GH, Peterson BR, Prabhu KS. Targeting of histone acetyltransferase p300 by cyclopentenone prostaglandin Δ(12)-PGJ(2) through covalent binding to Cys(1438). Chem Res Toxicol. 2012;25(2):337–47.
58. Hannafon BN, Carpenter KJ, Berry WL, Janknecht R, Dooley WC, Ding WQ. Exosome-mediated microRNA signaling from breast cancer cells is altered by the anti-angiogenesis agent docosahexaenoic acid (DHA). Mol Cancer. 2015;14:133.
59. Zheng Y, Li T, Luo X, Jia M, Han L, Yu H, et al. Sophora japonica extract inhibits ulcerative colitis and colitis-associated colon cancer via regulation of lipid metabolism. J Food Biochem. 2024;2024(1):5552747.
60. Huerta-Yépez S, Tirado-Rodriguez AB, Hankinson O. Role of diets rich in omega-3 and omega-6 in the development of cancer. Bol Med Hos Infant Mex. 2016;73(6):446–56.
61. Xu JC, Chen ZY, Huang XJ, Wu J, Huang H, Niu LF, et al. Multi-omics analysis reveals that linoleic acid metabolism is associated with variations of trained immunity induced by distinct BCG strains. Sci Adv. 2024;10(14):eadk8093.
62. Liu J, Mazzone PJ, Cata JP, Kurz A, Bauer M, Mascha EJ, et al. Serum free fatty acid biomarkers of lung cancer. Chest. 2014;146(3):670–9.
63. Zheng L, Hu F, Nie W, Lu J, Zhang B, Xu J, et al. The prognosis and metabolite changes of NSCLC patients receiving first-line immunotherapy combined chemotherapy in different M1c categories according to 9th edition of TNM classification. Cancer Med. 2024;13(17):e70223.
64. Li J, Liu K, Ji Z, Wang Y, Yin T, Long T, et al. Serum untargeted metabolomics reveal metabolic alteration of non-small cell lung cancer and refine disease detection. Cancer Sci. 2023;114(2):680–9.
65. Wang G, Qiu M, Xing X, Zhou J, Yao H, Li M, et al. Lung cancer scRNA-seq and lipidomics reveal aberrant lipid metabolism for early-stage diagnosis. Sci Transl Med. 2022;14(630):eabk2756.
66. Hagihara M, Kato H, Yamashita M, Shibata Y, Umemura T, Mori T, et al. Lung cancer progression alters lung and gut microbiomes and lipid metabolism. Heliyon. 2024;10(1): e23509.
67. Qu J, Xu S, Zhang Z, Chen G, Zhong Y, Liu L, et al. Evolutionary, structural and expression analysis of core genes involved in starch synthesis. Sci Rep. 2018;8(1):12736.
68. Yang Z, Zhuan B, Yan Y, Jiang S, Wang T. Integrated analyses of copy number variations and gene differential expression in lung squamous-cell carcinoma. Biol Res. 2015;48(1):47.
69. Lever J, Krzywinski M, Altman N. Principal component analysis. Nat Methods. 2017;14(7):641–2.
70. Tikkanen E, Jägerroos V, Holmes MV, Sattar N, Ala-Korpela M, Jousilahti P, et al. Metabolic biomarker discovery for risk of peripheral artery disease compared with coronary artery disease: lipoprotein and metabolite profiling of 31 657 individuals from 5 prospective cohorts. J Am Heart Assoc. 2021;10(23): e021995.

Publisher's Note