Association and biological pathways between lifetime occupational exposure to workplace hazards and incident chronic obstructive pulmonary disease and cardiovascular disease in middle-aged and older adults

Yang Yang[1,2#], Qichen Liu[1,3#], Filippos T. Filippidis[2], Peng Lu[4]*, Yuming Guo[5]*

[1] School of Public Health, Peking University, Beijing, China

[2] School of Public Health, Imperial College London, London, United Kingdom

[3] Institute for Environmental Health, Beijing Center for Disease Prevention and Control, Beijing, China

[4] School of Public Health, Binzhou Medical University, Yantai, Shandong, China

[5] School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia.

[#] These authors contributed equally as first authors

* Corresponding author to Professor Peng Lu, School of Public Health, Binzhou Medical University, Yantai, Shandong, China. E-mail: peng.lu@monash.edu.

Professor Yuming Guo, School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia. Email: yuming.guo@monash.edu.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The study is supported by the Taishan Scholar Foundation (No. tsqn202211228), the National Natural Science Foundation of China (No. 42307549).

Ethics approval and consent to participate

All participants provided informed consent through electronic signature at baseline assessment, and the UK Biobank study was approved by the North West-Haydock Research Ethics Committee (reference 21/NW/0157).

CRediT authorship contribution statement

Yang Yang: Conceptualization, Methodology, Formal analysis, Software, Writing original draft. Qichen Liu: Conceptualization, Methodology, Formal analysis, Data acquisition, Writing original draft. Filippos T. Filippidis: Conceptualization, Supervision, Methodology. Peng Lu: Conceptualization, Supervision, Project administration, Funding acquisition, Writing – Review & editing, Yuming Guo: Conceptualization, Supervision, Methodology, Project administration, Writing–Review & editing.

Acknowledgments

We are grateful to all participants and project teams in the UK Biobank study. This research has been conducted using the UK Biobank Resource under approved numbers 83151.

Data Availability Statement

The data of this study can be requested from the UK Biobank (https://www.ukbiobank.ac.uk/).

Abstract

The long-term impact of lifetime occupational exposure (LOE) on chronic obstructive pulmonary disease (COPD) and cardiovascular disease (CVD) risk remains unclear. This study examined associations between LOE and the risks of COPD and CVD in middle-aged and older adults. A prospective cohort study was conducted using UK Biobank data, including demographic, lifestyle, and genetic information. Cox proportional hazard models assessed associations of one-hazard (OLOE) and total-hazards LOE (TLOE) with cardiopulmonary outcomes. Mediation analyses explored the role of biomarkers and metabolites. Over a median 12.5-year follow-up, 2.4% (2,426/103,176) developed COPD and 20.6% (18,035/87,419) developed CVD. All OLOEs, except pesticide, were associated with elevated risks for both diseases. Higher TLOE was linked to increased COPD (HR: 1.21, 95% CI: 1.15–1.26) and CVD (HR: 1.05, 95% CI: 1.03–1.06) risks per exposure level increase. Clear dose-response relationships were observed. Inflammatory markers, such as white blood cell count, neutrophil count, and C-reactive protein, partially mediated these associations. Moreover, TLOE was significantly associated with the onset of a single cardiopulmonary disease and its progression to comorbidity. Our findings underscored the potential long-term cardiopulmonary burden of occupational hazards and supported the need for workplace hazard reduction to promote healthy aging.

Environmental Implication

This study reveals that lifetime occupational exposures to workplace hazards significantly increase long-term risks of COPD and CVD in aging populations, revealing an underrecognized environmental health challenge as populations age and working lives extend. By identifying inflammatory and metabolic pathways mediating these associations, our findings provide biological targets for early detection and preventive interventions in occupational settings. The demonstration that occupational exposures facilitate disease progression from single conditions to multimorbidity underscores the need for comprehensive workplace hazard elimination strategies throughout workers' careers rather than focusing solely on acute exposure prevention, and highlight the importance of integrating occupational health surveillance with chronic disease prevention programs to promote healthy aging and reduce the growing burden of cardiopulmonary multimorbidity.

Keywords Lifetime occupational exposure; Workplace hazards; Chronic obstructive pulmonary disease; Cardiovascular disease

Introduction

Chronic obstructive pulmonary disease (COPD) and cardiovascular diseases (CVDs) are leading causes of mortality and morbidity worldwide, collectively contributing to a substantial global health burden [1-4]. COPD affects over 210 million people worldwide with approximately 18.6 million deaths annually, and it is projected to become the third leading cause of death globally by 2030 [1,2]. CVDs account for an estimated 523 million prevalent cases and 18.6 million deaths each year, with both incidence and mortality projected to increase due to population aging and growth [3,4]. CVDs encompass a range of conditions affecting the heart and blood vessels, including coronary heart disease, cerebrovascular disease, peripheral arterial disease, and other vascular conditions [5]. Similarly, COPD, which encompasses conditions like emphysema and chronic bronchitis, is characterized by persistent airflow limitation and respiratory symptoms that progressively worsen over time [6]. Both conditions show increasing prevalence with age, disproportionately affecting middle-aged and older adults [2,4]. In the context of global demographic transitions toward aging societies and prolonged working lives, identifying modifiable risk factors is crucial for developing effective prevention strategies.

Occupational exposures represent a major yet under-investigated contributor to both COPD and CVDs risk. This gap is particularly significant as individuals spend about half their lives at work [7], where cumulative exposure to various workplace hazards may significantly impact health outcomes. Occupational exposures encompass diverse physical factors such as noise and extreme temperatures, as well as chemical factors including paints, asbestos, and pesticides [8-10]. They are among the most common in workplaces and are frequently studied together because they often co-occur in real-world settings. Recent evidence has begun to elucidate the associations between various occupational exposures and increased risk of cardiopulmonary diseases. For instance, a large-scale population-based cross-sectional analysis found that high exposure to airborne occupational pollutants, particularly vapors, gases, dusts, and fumes, was associated with an elevated risk of airflow obstruction, underscoring the relevance of occupational environments in the etiology of COPD [9]. In a cross-sectional study of over 7,000 Hispanic/Latino workers in the United States, occupational exposure to pesticides and metals was significantly associated with increased prevalence of CVD, including coronary heart disease and atrial fibrillation, after adjustment for potential covariates [10].

However, studies on occupational health impacts are often limited by small sample sizes, inadequate clinical parameter assessment, insufficient accounting for

confounding factors, and cross-sectional design. Notably, the assessment of occupational exposures presents significant methodological challenges due to their long-term and complex nature, often involving multiple substances at varying concentrations. Most previous studies assessing occupational exposures have relied on Job Exposure Matrices (JEMs), which assign exposure levels based on standardized occupational codes [11]. While JEMs provide a standardized approach that reduces recall bias and facilitates large-scale epidemiological analyses [9], they have notable limitations. First, JEMs assume uniform exposure levels within job categories, failing to capture individual-level variations in workplace conditions and account for temporal variations in exposure intensity (as many workers change jobs throughout their careers). Moreover, JEMs typically do not consider the cumulative burden of exposures across multiple jobs throughout an individual's career. Finally, JEMs are limited in their ability to reflect the complexity of real-world occupational exposures, where multiple hazardous factors often coexist and interact rather than acting in isolation in the workplace. For instance, an individual may experience multiple simultaneous exposures (e.g., exposed to chemical fumes, high noise levels, and extreme temperatures.) that may interact synergistically to affect cardiopulmonary health.

To overcome these limitations, our study employs a lifetime occupational exposure approach, which reconstructs individualized exposure profiles over the full course of a participant's occupational history. This method allows for more precise characterization of cumulative and co-exposure burdens. This approach is particularly valuable for middle-aged and older adults, who are more likely to have experienced prolonged and complex exposure patterns and are more physiologically susceptible to the long-term consequences of such exposures due to age-related decline in organ reserve and increased systemic vulnerability [12]. As populations worldwide continue to age and retirement ages extend, understanding these lifetime exposure patterns could better address chronic disease burden with natural aging processes.

Therefore, this study aims to investigate the associations between lifetime occupational exposure to workplace hazards and the development of COPD and CVD among middle-aged and older adults. Furthermore, we seek to explore the potential biological pathways linking occupational exposures to these cardiopulmonary outcomes, providing mechanistic insights that may inform preventive strategies and occupational health policy. Finally, we perform analyses within a disease-free cohort to comprehensively characterize the full spectrum of cardiopulmonary disease patterns attributable to lifetime occupational exposure.

Methods

Study Population

The UK Biobank (UKB) represented a large prospective cohort study, enrolling over half a million participants aged between 37 and 73 years during 2006-2010 throughout the United Kingdom. This study gathered extensive information both at the initial enrollment and during subsequent follow-up periods. At one of 22 assessment centers, participants engaged in both touchscreen and nurse-conducted questionnaires, underwent a series of physical measurements, and provided biological samples. The methodology and protocols of the UKB have been thoroughly outlined elsewhere [13]. The UKB study was granted under the approval of the North West Multicenter Research Ethical Committee (REC reference 21/NW/0157). All participants provided written informed consent. This research utilized the UK Biobank Resource under Application Number 83151.

In this study, we initially included 120,271 participants with complete lifetime occupational exposure data. However, participants younger than 45 years at baseline (N = 11,722) and those without a polygenic risk score (N = 2,496) were excluded. For COPD analysis, 2,877 participants with prevalent COPD were excluded, leaving 103,176 participants. For CVD analysis, 18,634 participants with prevalent CVD were excluded, resulting in 87,419 participants. Additionally, mediation analyses were conducted for each condition. Details of the selection process were presented in the Additional file 1: Figure S1.

Lifetime occupational hazard exposure

Participants provided a comprehensive lifetime occupational history, including all jobs and gap periods from the end of their education to the present. Each job was recorded chronologically, with details on the start and end years. For each job, participants answered ten questions regarding workplace environmental conditions, including exposure to noise, extreme temperatures, dust, chemical fumes, cigarette smoke, asbestos, paints, pesticides, and diesel exhaust. Responses were categorized as “Often,” “Sometimes,” “Rarely/Never,” or “Do not know,” with the latter treated as missing data. Each response was assigned a numerical value (1 for "Often," 0.5 for "Sometimes," and 0 for "Never"). Exposure scores were computed for each job both individually (for each specific hazard, range 0–1) and in aggregate (total hazard burden, range 0–10). Job duration (in years) was multiplied by average weekly working hours and by 52 (weeks per year) to convert exposure into total hours of exposure. Then, exposure scores were weighted by total hours of exposure to compute job-specific cumulative exposure score. Summing across all jobs yielded an individual's lifetime occupational exposure (LOE)

scores, calculated in two ways: for each individual hazard (OLOE, one-hazard lifetime occupational exposure) and for all hazards combined (TLOE, total lifetime occupational exposure). A detailed description of the LOE calculation is provided in Additional file 1: Supplementary Method. For the main analyses, OLOE and TLOE were categorized into four ordered exposure levels: "no exposure" and "low", "medium", and "high" exposure defined by tertiles among non-zero values. We also estimated hazard ratios per exposure-level increase to capture the monotonic trend.

Polygenic Risk Score

The detailed genotyping process, imputation, and quality control for the UKB have been described elsewhere [14]. Following the results of the prior genome-wide association study (GWAS) on COPD susceptibility, we incorporated 82 single nucleotide polymorphisms (SNPs) that demonstrated genome-wide significant associations with the disease [15]. Detailed information for each SNP and its effect size (β coefficient), is summarized in Additional file 1: Table S1. A polygenic risk score (PRS) for COPD was subsequently derived using the following formula: COPD PRS = (β1 × SNP1 + β2 × SNP2 + ···+βn × SNPn) *(N/sum of β-coefficient), where SNPn is the risk allele number of each SNP. In this study, PRS for CVD were obtained from the UKB PRS Release via the Research Access Platform. The PRS was generated using Bayesian analysis based on meta-analyses of summary statistics from external GWASs (standard PRS). The PRS calculation involved multiplying the genome-wide sum of the per-variant posterior effect size by allele frequencies. Detailed information regarding the methods can be found at https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=5202 and https://www.medrxiv.org/content/10.1101/2022.06.16.22276246v2. PRSs in the UKB Release have been shown to outperform a wide range of published PRSs and have been validated in an independent cohort (100,000 Genomes Project), demonstrating superior predictive accuracy and potential clinical utility [16].

Covariates

Age (continuous), sex (female or male), ethnicity (white or others), and assessment center (England, Wales, Scotland) were collected through touch-screen questionnaire. Body mass index (BMI) was calculated through weight divided by the square of height (kg/m^2). Education level was classified as “college or university degree”, “professional qualification”, “secondary school”, or “primary school”. Household income was categorized as “< £18000”, “£18000–30999”, “£31000–51999”, “£52000–100000”, and “> £100,000” per year”. The Townsend deprivation index (TDI, continuous) covered a wide range of information on social class, employment, and housing and

reflected the participant's area-based socioeconomic status (SES), with higher scores indicating greater deprivation [17]. Smoking status was categorized as "current", "former", or "never"; frequency of alcohol intake was measured as "≥ 3 times/week", "<3 times/week", or "never"; sleep duration was classified as "<7h per day", "7–8h per day", "> 8h per day". Physical activity was assessed in terms of Metabolic Equivalent Task (MET) minutes, derived from the short form of the International Physical Activity Questionnaire (IPAQ), and was subsequently classified into three levels: "high", "moderate", and "low" [18]. We defined a healthy diet score based on dietary priorities for cardiometabolic disease [19], where a higher score indicated healthier dietary habits. Definitions of each component of a healthy diet score were described in the Additional file 1: Table S2.

Mediators

Blood sample collection and processing in the UK Biobank was conducted under strict protocols with participants' informed consent during baseline recruitment. Approximately 4 ml of blood was drawn from each participant and fractionated into its components. The samples were preserved at −80 °C and underwent analysis within 24 hours using the Beckman Coulter LH750 instrument (https://biobank.ndph.ox.ac.uk/ukb/ukb/docs/haematology.pdf). To ensure reliability, blood biomarkers undergo rigorous quality control measures and external validation, providing comprehensive information on assay performance (https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_biochemistry.pdf). Plasma samples from about one-fifth of participants in the baseline recruitment were tested for metabolites using Nightingale Health's NMR-based high-throughput metabolic biomarker analysis platform [20]. According to the online documentation provided by the UKB (https://biobank.ndph.ox.ac.uk/ukb/ukb/docs/nmrm_companion_doc.pdf), a total of 168 metabolites could be utilized directly for epidemiological analyses without requiring any preprocessing, with most reflecting lipid metabolism and a few showing protein hydrolysis, glycolysis, and ketolysis.

As this study aimed to explore potential biological pathways, we included a comprehensive set of 36 blood biomarkers, including those related to inflammation, erythrocyte function, liver and renal function, and circulating metabolic biomarkers, along with 168 plasma metabolites, were selected as potential mediators. Inflammation-related biomarkers included white blood cell (WBC) count, neutrophil count, neutrophil percentage, monocyte count, monocyte percentage, lymphocyte count, lymphocyte percentage, basophill count, basophill percentage, C reactive protein, and platelet count.

Erythrocyte-related biomarkers included erythrocyte count, high light scatter reticulocyte count, reticulocyte count, red blood cell distribution width, haematocrit percentage, and haemoglobin concentration. Renal function-related biomarkers included cystatin C, urate, urea and creatinine. Liver function-related biomarkers included alanine aminotransferase, alkaline phosphatase, aspartate aminotransferase, gamma-glutamyl transferase, total bilirubin, total protein, and albumin. Circulating metabolic biomarkers included glycated hemoglobin, glucose, cholesterol, triglycerides, low-density lipoprotein, high-density lipoprotein, apolipoprotein A, and apolipoprotein B. The mean concentration for blood biomarkers and plasma metabolites can be seen in Additional file 2: Table S1, Table S8.

Outcome

The primary outcomes of interest were incident COPD and CVD. These outcomes were ascertained by linkage to health-related records, including hospital inpatient, death register, and primary care. Details on disease coding are provided in the Additional file 1: Table S18. Prevalent status was defined based on self-report doctor diagnosed conditions and health-related records at baseline visit. Follow-up time was calculated from the date of the baseline assessment to the earliest of the following: date of first recorded COPD or CVD diagnosis, death, loss to follow-up, or end of study (30 September 2021 for England and Wales; 31 October 2021 for Scotland).

Statistical analysis

Descriptive characteristics of the included participants are presented as frequencies with percentages for categorical variables and mean with standard deviations (SD) for continuous variables. Missing data of covariates were imputed via the R package 'missForest' [21].

To assess the effects of OLOE and TLOE on COPD and CVD, Cox proportional hazard regression models with follow-up time as the underlying timescale were conducted. The proportional hazards assumption was examined through statistical tests based on scaled Schoenfeld residuals. Assessment center was violation of the assumption and corrected through stratifying. Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry. Also, we estimated hazard ratios (HRs) per exposure level increase for TLOE and calculated P for trend to assess the significance of the trend. We further explored the dose-response relationship between TLOE and cardiopulmonary outcomes, using restricted cubic

splines (RCS) with knots placed at the 5th, 35th, 65th, and 95th percentiles of each exposure (reference was the 5th percentile).

All biomarkers and metabolites were standardized (z-score) before entering formal analysis and TLOE was analyzed per exposure level increase. The mediators were selected by the following analyses [22]. First, multiple linear regression models were applied to evaluate the association of TLOE with biomarkers/metabolites. Second, adjusted Cox regression models adjusting for were implemented to examine the relationship between biomarkers/metabolites and incident cardiopulmonary outcomes (COPD, CVD, respectively). All p-values in the regression models were corrected by the Benjamini-Hochberg false discovery rate (FDR) for multiple comparisons and was considered as statistically significant when an FDR-adjusted p value (P_{FDR}-value) less than 0.05. Then, all biomarkers and metabolites were subsequently applied mediation analyses via the "CMAverse" package [23], where the direct, indirect, and total effects on the hazard ratio, as well as the proportion mediated (PM) were reported. The parametric bootstrapping method (n=1000 times) was used to calculate 95% confidence intervals (CIs) and p-values. The biomarkers and metabolites that exhibited significance in the aforementioned steps were finally presented as potential mediators and a multi-mediator model was used to assess the combined mediating effect of all selected mediators.

To comprehensively assess the full spectrum of cardiopulmonary disease patterns induced by TLOE, we conducted analyses in a disease-free cohort, excluding all participants with prevalent COPD or CVD at baseline. We assessed associations between TLOE and four outcomes: incident COPD, incident CVD, the first occurrence of either disease, and their comorbidity.

Additionally, we conducted a stratification analysis to explore whether the association between TLOE and cardiopulmonary outcomes (COPD and CVD respectively) differed across various subgroups: age (<60 years and ≥60 years), sex (female and male), PRS (≤ median and > median), education level (college/university and other), household income (≥ £52,000 and < £52,000 per year), Townsend Deprivation Index (≤ median and > median), healthy diet score (≤ median and > median), physical activity (high and other), drinking status (never and other), smoking status (never and other), and sleep duration (7-8 hours/day and other).

In the supplementary analyses, we explored associations between TLOE and several major CVD subtypes, including coronary heart disease, atrial fibrillation, stroke, and

heart failure. To further examine the interplay between COPD and CVD, we conducted bidirectional analyses evaluating the associations of TLOE and prevalent CVD at baseline with the risk of incident COPD, and the associations of TLOE and prevalent COPD at baseline with the risk of incident CVD. Additionally, we evaluated the interaction between TLOE and genetic susceptibility (COPD PRS, CVD PRS, respectively) on the risk of incident COPD and CVD. Interactions were evaluated using likelihood ratio tests to compare models with and without a cross-product term. Moreover, to calculate the cumulative incidence of cardiopulmonary outcomes based on TLOE, we fitted the Fine–Gray proportional sub-distribution hazards model and applied the cumulative incidence function [24].

A series of sensitivity analyses were performed to assess the robustness of the findings and to address potential biases. First, to evaluate the impact of job duration measurement precision, we applied a 0.5-year adjustment to all reported job durations, replacing the 1-year adjustment used in the main analysis for jobs reported with identical start and end years. Second, cases occurring within the first two years were excluded to reduce the potential influence of reverse causality. Third, for COPD outcomes, we conducted two additional analyses to address potential misclassification: (1) excluding participants with baseline $FEV_1/FVC < 0.70$ and (2) excluding participants who developed asthma during follow-up. Fourth, compared to the main analysis where missing values were imputed using the 'missForest' package, missing values for continuous covariates were addressed by using the median values, and missing values for categorical covariates were treated as a separate category. Fifth, we further adjusted for prevalent depression, anxiety, asthma, hypertension, diabetes, COPD (for incident CVD) and CVD (for incident COPD) at baseline. Sixth, we further adjusted for medication use at baseline, including aspirin, cholesterol-lowering drugs, antihypertensive agents, and insulin. Seventh, we additionally adjusted for family history of diabetes, cardiovascular disease, hypertension, and lung disease. Eighth, we further accounted for environmental exposures, including ambient air pollutants (NO_2, NO_x, $PM_{2.5}$, $PM_{2.5-10}$, PM_{10}), use of solid fuels for heating or cooking, and exposure to passive smoking. Finally, we performed multistate analyses to evaluate associations between TLOE and transitions from baseline to first cardiopulmonary disease and subsequently to multimorbidity.

All statistical analyses were conducted using R software version 4.3.1. All statistical tests were two-sided, and we considered a *P* value of less than 0.05 to be statistically significant.

Results

Characteristics of the study participants and workplace hazards

The baseline characteristics of the study participants are presented in Table 1. The study population size varied by outcome due to the exclusion of prevalent cases at baseline. During a median follow-up of 12.5 years, 2,426 (2.4%) of 103,176 participants developed COPD and 18,035 (20.6%) of 87,419 participants developed CVD. Characteristics of workplace hazard exposure are presented in Figure 1. The proportion of zero-proportion ranged from 36.4% in noisy to 95.9% in pesticides for COPD and from 36.7% in noisy to 96.0% in pesticides for CVD. Positive correlations were observed between various occupational hazards, with Spearman correlation coefficients ranging from 0.07 to 0.62 in both COPD and CVD.

Higher OLOE associated with elevated COPD and CVD risks

Figure 2 presents the associations between each OLOE and the risks of COPD and CVD. For COPD, high exposure to dusty environments was associated with a 49% increased risk (HR = 1.49, 95% CI: 1.34–1.67), with elevated risks observed for chemical fumes (HR = 1.38, 95% CI: 1.22–1.56), smoke (HR = 1.50, 95% CI: 1.35–1.67), asbestos (HR = 1.22, 95% CI: 1.03–1.45), paints (HR = 1.19, 95% CI: 1.04–1.37), diesel exhaust (HR = 1.34, 95% CI: 1.16–1.54), noise (HR = 1.47, 95% CI: 1.31–1.64), cold (HR = 1.27, 95% CI: 1.13–1.41), and hot (HR = 1.29, 95% CI: 1.15–1.44). For CVD, high exposure to paints (HR = 1.20, 95% CI: 1.12–1.27), asbestos (HR = 1.12, 95% CI: 1.05–1.21), diesel exhaust (HR = 1.13, 95% CI: 1.07–1.20), noise (HR = 1.10, 95% CI: 1.06–1.15), cold (HR = 1.12, 95% CI: 1.08–1.17), and hot (HR = 1.11, 95% CI: 1.06–1.15) was associated with increased cardiovascular risk, although the magnitude of association was less pronounced compared with COPD. Exposure to pesticides was not significantly associated with either COPD or CVD.

Higher TLOE associated with elevated COPD and CVD risks

Associations of TLOE and the risks of COPD and CVD are presented in Figure 3A. For COPD and CVD, after adjustment for potential covariates, individuals with high exposure had significantly increased risks compared to those with no exposure, with all trend analyses yielding strong statistical significance (P for trend < 0.001). Per exposure level increase, the adjusted HRs were 1.21 (95% CI: 1.15–1.26) for COPD and 1.05 (95% CI: 1.03–1.06) for CVD. Figure 3B illustrates the dose–response relationship between TLOE and cardiopulmonary outcomes. Dose–response curves generated using restricted cubic splines showed steep and nonlinear trajectories (P for nonlinearity < 0.05): HRs increased progressively with higher TLOE values both for COPD and CVD, indicating a clear and positive association.

Inflammatory biomarkers partially mediating the associations

Associations of biomarkers and metabolites with TLOE and cardiopulmonary outcomes are presented in Figure 4. Several inflammatory biomarkers showed significant associations with both TLOE and cardiopulmonary outcomes. In contrast, although multiple metabolites were associated with cardiopulmonary outcomes, no significant associations were identified between metabolites and TLOE. Therefore, only biomarkers were included in the subsequent mediation analysis. Biomarkers mediated the association between TLOE and cardiopulmonary outcomes as shown in Figure 5. Inflammatory markers significantly mediated these relationships (FDR < 0.05), suggesting that TLOE may increase cardiopulmonary risk through inflammation-related pathways. Specifically, white blood cell count (WBC), neutrophil count, and C-reactive protein (CRP) accounted for 0.75%, 1.70%, and 0.74% of the TLOE–COPD association, and 1.20%, 1.74%, and 1.78% of the TLOE–CVD association, respectively. In the combined mediation analysis, these three inflammatory markers together explained 2.20% of the TLOE–COPD association and 3.00% of the TLOE–CVD association. However, no metabolites are available as potential mediators of the association between TLOE and incident COPD and CVD, respectively. Detailed information on the associations between TLOE and biomarkers, and between biomarkers and outcomes, is presented in Additional file 2: Tables S2–S7. Detailed information on the associations between TLOE and metabolites, and between metabolites and outcomes, is provided in Additional file 2: Tables S9–S12.

Higher TLOE associated with elevated risk of cardiopulmonary comorbidity

After excluding participants with baseline COPD or CVD, TLOE demonstrated significant associations with all cardiopulmonary outcomes (Table 2). High exposure was associated with increased risks of incident COPD, CVD, any cardiopulmonary disease, and COPD-CVD comorbidity. Per exposure level increase, the adjusted HRs were 1.20 (95% CI: 1.14–1.27) for COPD, 1.04 (95% CI: 1.03–1.06) for CVD, 1.05 (95% CI: 1.04–1.07) for any disease, and 1.15 (95% CI: 1.06–1.24) for comorbidity. All trend analyses yielded strong statistical significance (P for trend < 0.001).

Sensitivity analysis

Cumulative risk of COPD and CVD stratified by TLOE were presented in Additional file 1: Figures S2-S3. Joint associations of TLOE and PRS on these cardiopulmonary outcomes were displayed in Additional file 1: Figures S4-S5. As shown in Additional File 1: Tables S3–S4, the associations between TLOE and cardiorespiratory outcomes were consistent across various factors, with progressively higher risks observed as

TLOE categories increased. Associations of TLOE and the risks of CHD, AF, HF and stroke are shown in Additional File 1: Tables S5. Results regarding the combined effects of TLOE and prevalent CVD on incident COPD, as well as the combined effects of TLOE and prevalent COPD on incident CVD, are shown in Additional File 1: Tables S6–S7, with significant interactions observed. The primary findings remained robust when using 0.5-year adjustment for job duration (Additional file 1: Table S8, Figure S6). Sensitivity analyses confirmed the robustness of the results after removing cases within the first two years of follow-up, excluding individuals with $FEV_1/FVC < 0.70$ at baseline, excluding incident asthma cases during follow-up, and using covariates without imputation (Additional file 1: Tables S9–S12). Similarly, additional adjustments for prevalent diseases, medication use, family history, and environmental factors (Additional file 1: Tables S13–S16) yielded consistent results. Multistate analyses showed that TLOE was significantly associated with transition from baseline to first cardiopulmonary disease and from first cardiopulmonary disease to comorbidity (Additional file 1: Table S17).

Discussion

This study provides compelling evidence for the associations between lifetime occupational exposure to workplace hazards and elevated risks of COPD and CVD among middle-aged and older adults. Except for pesticides, all OLOEs showed significant associations with higher risks of COPD and CVD across increasing exposure levels. We observed clear dose-response relationships, with higher levels of TLOE corresponding to increased HRs for all outcomes. Notably, our investigation into potential mediating pathways identified three key inflammatory biomarkers (white blood cell count, neutrophil count, and C-reactive protein) that partially mediated the relationship between TLOE and cardiopulmonary outcomes, while metabolite pathways showed no significant mediation effects.

Previous research has documented associations between specific occupational exposures and adverse cardiopulmonary outcomes. For example, high exposure to vapors, gases, dusts, and fumes has been linked to increased risk of airflow obstruction [9], while exposure to pesticides and metals has been associated with higher prevalence of cardiovascular conditions [10]. Consistent with these findings, we observed that most OLOE scores were significantly associated with increased risks of COPD and CVD. However, no statistically significant associations were found between pesticide-related OLOE and either outcome, possibly due to the high proportion (96%) of zero-exposure in our sample, which substantially limited exposure variability and statistical power. Our study addressed a key limitation of previous research by incorporating a

life-course and multi-exposure perspective, and further assessed the association between TLOE, which reflects the burden of all occupational hazards, and cardiopulmonary outcomes. Our dose–response analyses revealed a steeper gradient of hazard ratios with increasing TLOE for both COPD and CVD, without an apparent threshold effect, highlighting the importance of minimizing occupational exposures even at low levels.

Traditional methods of occupational exposure assessment, such as job-exposure matrices (JEMs), rely on standardized exposure levels for specific occupations and often assume uniform exposure within job categories [11]. This approach fails to account for individual variability, temporal dynamics, and the interactive effects of co-occurring hazards [25]. In contrast, the LOE framework enables a more comprehensive evaluation of cumulative occupational exposures across the working life by explicitly integrating exposure intensity, duration, and sequence into a single, life-course–oriented construct, rather than treating exposure as a static attribute of job title alone. Beyond simply reclassifying exposure categories, the LOE model reconstructs individual exposure trajectories over time, allowing differentiation between workers with the same occupation but divergent patterns of task demands, work schedules, and co-exposure profiles. The current occupational safety standards primarily focus on preventing acute injuries rather than addressing the long-term consequences of chronic, low-level exposures [26,27]. Moreover, they typically establish permissible limits for individual hazards while ignoring multifactorial exposures and failing to consider their combined biological impact [28,29]. This limitation underscores the need for a paradigm shift toward cumulative risk models. A more holistic regulatory approach that explicitly considers the interaction of multiple and long-term exposures is essential for safeguarding worker health and addressing the underrecognized risks associated with chronic occupational exposures.

The consistency of these associations across various subgroup analyses has important implications for public health and occupational medicine. The persistence of associations across age groups (<60 years and ≥60 years) suggests that the cardiovascular impacts of occupational exposures affect individuals both during active working years and continue into later life. This highlights the need for early intervention and ongoing monitoring of workplace exposures. Sex-stratified analysis revealed consistent effects, with more pronounced in men. This may reflect sex-specific biological vulnerabilities and physiological responses or differences in exposure intensity within similar job categories [30]. Interestingly, while prior research has established that higher socioeconomic status, healthier lifestyle behaviors, and lower

genetic susceptibility are each associated with reduced cardiopulmonary risk [15, 31-35], our findings demonstrate that these advantages do not fully offset the adverse effects of occupational exposures. The persistence of associations in these typically lower-risk subgroups suggests that occupational hazards operate through biological pathways that circumvent conventional protective mechanisms, representing an independent risk domain that transcends traditional sociodemographic and behavioral determinants. These findings therefore provide empirical support for a primary prevention agenda that prioritizes structural changes to the work environment, such as stricter exposure limits, comprehensive surveillance of mixed exposures, enforcement of control technologies, and incentive structures that reward employers for sustained exposure reduction, as essential complements to traditional lifestyle-based prevention strategies. These results also suggest that meaningful reductions in cardiopulmonary burden will require shifting prevention efforts “upstream” to policies and regulations that systematically reduce hazardous occupational exposures, rather than expecting downstream, individual-level risk factor management to compensate for avoidable risks at work.

We observed significant interactions between TLOE and genetic susceptibility on incident COPD and CVD in supplementary analysis. These interactions suggest that individuals with higher genetic susceptibility may be particularly vulnerable to the cardiopulmonary effects of workplace exposures, reinforcing the importance of gene–environment interplay in the development of chronic disease. Importantly, even individuals with lower PRS showed an increased cardiopulmonary risk with higher occupational exposures, indicating that workplace interventions could benefit a broad population, regardless of genetic predisposition. Our findings align with the "dual burden" hypothesis, which posits that both genetic and environmental factors work together to amplify disease risk. From a practical perspective, these findings suggest that integrating genetic susceptibility into occupational surveillance may help identify groups who could benefit most from targeted prevention, although the routine clinical use of polygenic risk scores in occupational health is not yet established and will require further validation and ethical consideration. While genetic risk cannot be modified, reducing occupational exposures offers a tangible and universally applicable strategy for mitigating cardiopulmonary risk across all genetic risk strata. From a public health perspective, primary prevention through workplace exposure reduction remains the most equitable and feasible approach, as it protects all workers without requiring genetic stratification.

The observed associations between TLOE and cardiopulmonary outcomes likely

operate through interconnected biological mechanisms. Our mediation analysis used 36 blood biomarkers and 168 plasma metabolites measured at baseline, prior to disease onset, and identified inflammatory markers (white blood cell count, neutrophil count, and C-reactive protein) as the only statistically significant mediators linking lifetime occupational exposure to subsequent cardiopulmonary outcomes. First, the oxidative stress pathway is initiated when inhaled occupational toxicants generate excessive reactive oxygen species (ROS). These ROS activate redox-sensitive transcription factors, including NF-κB and AP-1, which induce the expression of proinflammatory cytokines such as interleukin-6 (IL-6) and tumor necrosis factor-alpha (TNF-α) [36]. The resulting systemic inflammation is characterized by elevated white blood cell (WBC) and neutrophil counts, along with increased hepatic synthesis of C-reactive protein (CRP). These inflammatory mediators contribute to alveolar wall destruction and airway remodeling in COPD, and to endothelial dysfunction, lipid oxidation, and atherogenesis in CVD [37-39]. Second, the innate immune activation pathway begins when pattern recognition receptors (PRRs) on airway epithelial cells, including Toll-like receptors (TLRs) and nucleotide-binding oligomerization domain-like receptors (NOD-like receptors or NLRs), recognize inhaled occupational particles as danger signals [40,41]. This recognition activates the NOD-like receptor family pyrin domain-containing 3 (NLRP3) inflammasome, leading to the release of interleukin-1β (IL-1β) [40-42]. IL-1β promotes leukocyte proliferation, neutrophil recruitment, and hepatic CRP production, thereby amplifying systemic inflammation [40-42]. In the lungs, this drives neutrophil-mediated proteolytic damage, mucus hypersecretion, and emphysematous remodeling [43]. In the vasculature, IL-1β increases the expression of vascular adhesion molecules such as VCAM-1 and ICAM-1, facilitating leukocyte–endothelium interactions and promoting atherosclerotic plaque formation and instability [44]. Third, physical occupational stressors including heat, cold, and noise can activate the sympathetic nervous system (SNS) and the renin-angiotensin system (RAS), leading to neuroendocrine responses that enhance bone marrow output of leukocytes, including neutrophils, and increase systemic CRP levels [45]. These responses contribute to pulmonary hypertension and vascular remodeling in COPD, and accelerate arterial stiffening, myocardial hypertrophy, and heart failure progression in CVD [46,47].

After excluding participants with pre-existing COPD or CVD at baseline, we found that TLOE was significantly associated not only with the incidence of COPD and CVD as individual outcomes, but also with their co-occurrence. Multistate analysis further revealed that TLOE was associated with an elevated risk of transition from a disease-free state to the onset of a single cardiopulmonary condition, as well as from a single

condition to subsequent comorbidity. These findings highlight potential interactions between disease pathways and emphasize the importance of integrating occupational health considerations into the prevention and management of chronic cardiopulmonary diseases. Our supplementary analysis revealed a critical bidirectional interaction between CVD and COPD in the context of occupational hazard exposure. The presence of baseline CVD substantially amplifies the deleterious effects of occupational exposures on subsequent COPD development, with a similar pattern observed for COPD and CVD risk. This interaction highlights that current occupational exposure limits, established based on healthy worker populations, may overlook the heightened vulnerability of individuals with pre-existing cardiopulmonary conditions. Clinically, these findings mandate integration of detailed occupational histories into cardiopulmonary risk stratification and suggest that patients with cardiorespiratory conditions require enhanced workplace protections and closer surveillance. For public health policy, our results provide the epidemiological foundation for developing risk-stratified occupational safety protocols where exposure limits are calibrated to individual vulnerability profiles rather than population averages, representing a paradigm shift toward precision occupational medicine and potentially preventing substantial cardiopulmonary morbidity in working populations worldwide.

Our study has several strengths. First, it benefits from a large sample size, long follow-up period, and detailed data on demographics, socioeconomic status, lifestyle factors, genetics, and objective medical information. Second, the use of the LOE model for occupational exposure assessment represents a novel and rigorous methodological approach. This model captures the cumulative nature of workplace hazards throughout an individual's career, offering a more comprehensive and accurate reflection of long-term cardiopulmonary risk compared to traditional JEM. Third, the integration of inflammatory biomarkers as mediators in the exposure-disease relationship provides valuable mechanistic insights, enhancing our understanding of the biological pathways linking occupational exposures to cardiopulmonary outcomes.

However, several limitations of this study should be acknowledged. First, due to the observational nature of the design, causal inferences cannot be definitively established, and although we adjusted for a wide range of potential confounders, the possibility of residual confounding cannot be ruled out. Second, relying on medical records may have led to underdiagnosis of chronic conditions such as COPD and CVD, particularly in individuals who did not seek medical care or were never formally diagnosed. To address potential misclassification of COPD, we conducted sensitivity analyses by excluding participants with FEV1/FVC < 0.70 at baseline and those who developed asthma during

follow-up, and both analyses yielded consistent results, supporting the robustness of our findings. Third, the reliance on self-reported data for occupational history may introduce the potential for recall and reporting bias, which could affect the accuracy of exposure assessments, and lead to non-differential misclassification. As noted in prior studies, non-differential misclassification of exposure could result in an underestimation of the true association, potentially biasing the results toward the null [48]. Fourth, lifetime occupational exposure was assessed once at baseline and was not updated during follow-up. Because most participants were middle-aged or older at recruitment, the baseline assessment already captured the majority of cumulative exposure accumulated during earlier and mid-career years, when exposure intensity is usually highest. Although younger participants may have acquired additional exposure after baseline, such post-baseline exposure would introduce non-differential measurement error, which typically biases results toward the null. Nevertheless, stratified analyses by age yielded consistent results, suggesting that this limitation may have had a limited impact on the overall findings. Fifth, although we included a broad panel of biomarkers and metabolites, these measures may still not fully capture all biological pathways through which occupational exposures influence cardiopulmonary risk. Finally, this cohort predominantly included individuals of European descent (mostly White British), which may limit the generalizability of the findings to other ethnic groups, such as Asians and Blacks. The UKB aimed to be representative of the general population, but it is unrepresentative in terms of lifestyle due to a healthy volunteer selection bias [49]. Moreover, UKB participants are less likely to come from high-exposure manual or industrial occupations, which may lead to an underrepresentation of individuals with the greatest workplace hazard burden. These characteristics may limit the generalizability of absolute exposure levels and suggest a need for replication in more diverse and occupationally heterogeneous populations to confirm the broader applicability of the results.

Conclusion

In conclusion, higher lifetime occupational exposure was associated with increased risks of incident COPD and CVD. Importantly, these risks extended beyond single diseases to cardiopulmonary comorbidity, highlighting cumulative workplace exposure as a distinct and long-term determinant of cardiopulmonary health in aging populations. These results advocate for integrating occupational exposure history into routine clinical risk assessments to facilitate early detection and targeted prevention strategies.

Reference

1. Safiri S, Carson-Chahhoud K, Noori M, Nejadghaderi SA, Sullman MJM, Ahmadian Heris J, Ansarin K, Mansournia MA, Collins GS, Kolahi AA, Kaufman JS. Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990-2019: results from the Global Burden of Disease Study 2019. BMJ. 2022 Jul 27;378:e069679. doi: 10.1136/bmj-2021-069679. PMID: 35896191; PMCID: PMC9326843.
2. Wang Z, Lin J, Liang L, Huang F, Yao X, Peng K, Gao Y, Zheng J. Global, regional, and national burden of chronic obstructive pulmonary disease and its attributable risk factors from 1990 to 2021: an analysis for the Global Burden of Disease Study 2021. Respir Res. 2025 Jan 2;26(1):2. doi: 10.1186/s12931-024-03051-2. PMID: 39748260; PMCID: PMC11697803.
3. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton AZ, Benjamin EJ, Benziger CP, Bonny A, Brauer M, Brodmann M, Cahill TJ, Carapetis J, Catapano AL, Chugh SS, Cooper LT, Coresh J, Criqui M, DeCleene N, Eagle KA, Emmons-Bell S, Feigin VL, Fernández-Solà J, Fowkes G, Gakidou E, Grundy SM, He FJ, Howard G, Hu F, Inker L, Karthikeyan G, Kassebaum N, Koroshetz W, Lavie C, Lloyd-Jones D, Lu HS, Mirijello A, Temesgen AM, Mokdad A, Moran AE, Muntner P, Narula J, Neal B, Ntsekhe M, Moraes de Oliveira G, Otto C, Owolabi M, Pratt M, Rajagopalan S, Reitsma M, Ribeiro ALP, Rigotti N, Rodgers A, Sable C, Shakil S, Sliwa-Hahnle K, Stark B, Sundström J, Timpel P, Tleyjeh IM, Valgimigli M, Vos T, Whelton PK, Yacoub M, Zuhlke L, Murray C, Fuster V; GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group. Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. J Am Coll Cardiol. 2020 Dec 22;76(25):2982-3021. doi: 10.1016/j.jacc.2020.11.010. Erratum in: J Am Coll Cardiol. 2021 Apr 20;77(15):1958-1959. doi: 10.1016/j.jacc.2021.02.039. PMID: 33309175; PMCID: PMC7755038.
4. Qu C, Liao S, Zhang J, Cao H, Zhang H, Zhang N, Yan L, Cui G, Luo P, Zhang Q, Cheng Q. Burden of cardiovascular disease among elderly: based on the Global Burden of Disease Study 2019. Eur Heart J Qual Care Clin Outcomes. 2024 Mar 1;10(2):143-153. doi: 10.1093/ehjqcco/qcad033. PMID: 37296238; PMCID: PMC10904724.
5. Mensah GA, Roth GA, Fuster V. The Global Burden of Cardiovascular Diseases and Risk Factors: 2020 and Beyond. J Am Coll Cardiol. 2019 Nov 19;74(20):2529-2532. doi: 10.1016/j.jacc.2019.10.009. PMID: 31727292.
6. Kim V, Criner GJ. Chronic bronchitis and chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2013 Feb 1;187(3):228-37. doi: 10.1164/rccm.201210-1843CI. Epub 2012 Nov 29. PMID: 23204254; PMCID: PMC4951627.
7. Houot MT, Homère J, Goulard H, Garras L, Delabre L, Pilorget C. Lifetime occupational exposure proportion estimation methods: a sensitivity analysis in the general population. Int Arch Occup Environ Health. 2021 Oct;94(7):1537-1547. doi: 10.1007/s00420-021-01691-1. Epub 2021 Apr 13. PMID: 33847787.

8. Miguet M, Rukh G, Titova OE, Schiöth HB. Important Difference between Occupational Hazard Exposure among Shift Workers and Other Workers; Comparing Workplace before and after 1980. Int J Environ Res Public Health. 2020 Oct 15;17(20):7495. doi: 10.3390/ijerph17207495. PMID: 33076288; PMCID: PMC7602472.
9. Sadhra SS, Mohammed N, Kurmi OP, Fishwick D, De Matteis S, Hutchings S, Jarvis D, Ayres JG, Rushton L. Occupational exposure to inhaled pollutants and risk of airflow obstruction: a large UK population-based UK Biobank cohort. Thorax. 2020 Jun;75(6):468-475. doi: 10.1136/thoraxjnl-2019-213407. Epub 2020 May 6. PMID: 32376731.
10. Bulka CM, Daviglus ML, Persky VW, Durazo-Arvizu RA, Lash JP, Elfassy T, Lee DJ, Ramos AR, Tarraf W, Argos M. Association of occupational exposures with cardiovascular disease among US Hispanics/Latinos. Heart. 2019 Mar;105(6):439-448. doi: 10.1136/heartjnl-2018-313463. Epub 2018 Dec 11. PMID: 30538094; PMCID: PMC6580877.
11. Sadhra S, Kurmi OP, Sadhra SS, Lam KB, Ayres JG. Occupational COPD and job exposure matrices: a systematic review and meta-analysis. Int J Chron Obstruct Pulmon Dis. 2017 Feb 22;12:725-734. doi: 10.2147/COPD.S125980. PMID: 28260879; PMCID: PMC5327910.
12. Sundstrup E, Hansen ÅM, Mortensen EL, Poulsen OM, Clausen T, Rugulies R, Møller A, Andersen LL. Cumulative occupational mechanical exposures during working life and risk of sickness absence and disability pension: prospective cohort study. Scand J Work Environ Health. 2017 Sep 1;43(5):415-425. doi: 10.5271/sjweh.3663. Epub 2017 Aug 7. PMID: 28783203.
13. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R. 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 Mar 31;12(3):e1001779. doi: 10.1371/journal.pmed.1001779. PMID: 25826379; PMCID: PMC4380465.
14. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018 Oct;562(7726):203-209. doi: 10.1038/s41586-018-0579-z. Epub 2018 Oct 10. PMID: 30305743; PMCID: PMC6786975.
15. Sakornsakolpat P, Prokopenko D, Lamontagne M, Reeve NF, Guyatt AL, Jackson VE, Shrine N, Qiao D, Bartz TM, Kim DK, Lee MK, Latourelle JC, Li X, Morrow JD, Obeidat M, Wyss AB, Bakke P, Barr RG, Beaty TH, Belinsky SA, Brusselle GG, Crapo JD, de Jong K, DeMeo DL, Fingerlin TE, Gharib SA, Gulsvik A, Hall IP, Hokanson JE, Kim WJ, Lomas DA, London SJ, Meyers DA, O'Connor GT,

Rennard SI, Schwartz DA, Sliwinski P, Sparrow D, Strachan DP, Tal-Singer R, Tesfaigzi Y, Vestbo J, Vonk JM, Yim JJ, Zhou X, Bossé Y, Manichaikul A, Lahousse L, Silverman EK, Boezen HM, Wain LV, Tobin MD, Hobbs BD, Cho MH; SpiroMeta Consortium; International COPD Genetics Consortium. Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. Nat Genet. 2019 Mar;51(3):494-505. doi: 10.1038/s41588-018-0342-2. Epub 2019 Feb 25. PMID: 30804561; PMCID: PMC6546635.

16. Thompson DJ, Wells D, Selzam S, Peneva I, Moore R, Sharp K, Tarran WA, Beard EJ, Riveros-Mckay F, Giner-Delgado C, Palmer D, Seth P, Harrison J, Futema M; Genomics England Research Consortium; McVean G, Plagnol V, Donnelly P, Weale ME. A systematic evaluation of the performance and properties of the UK Biobank Polygenic Risk Score (PRS) Release. PLoS One. 2024 Sep 18;19(9):e0307270. doi: 10.1371/journal.pone.0307270. PMID: 39292644; PMCID: PMC11410272.
17. Foster HME, Celis-Morales CA, Nicholl BI, Petermann-Rocha F, Pell JP, Gill JMR, O'Donnell CA, Mair FS. The effect of socioeconomic deprivation on the association between an extended measurement of unhealthy lifestyle factors and health outcomes: a prospective analysis of the UK Biobank cohort. Lancet Public Health. 2018 Dec;3(12):e576-e585. doi: 10.1016/S2468-2667(18)30200-7. Epub 2018 Nov 20. PMID: 30467019.
18. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, Sallis JF, Oja P. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc. 2003 Aug;35(8):1381-95. doi: 10.1249/01.MSS.0000078924.61453.FB. PMID: 12900694.
19. Said MA, Verweij N, van der Harst P. Associations of Combined Genetic and Lifestyle Risks With Incident Cardiovascular Disease and Diabetes in the UK Biobank Study. JAMA Cardiol. 2018 Aug 1;3(8):693-702. doi: 10.1001/jamacardio.2018.1717. PMID: 29955826; PMCID: PMC6143077.
20. 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 Feb;8(1):192-206. doi: 10.1161/CIRCGENETICS.114.000216. PMID: 25691689.
21. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012 Jan 1;28(1):112-8. doi: 10.1093/bioinformatics/btr597. Epub 2011 Oct 28. PMID: 22039212.
22. Hu W, Liu BP, Jia CX. Association and biological pathways between lung function and incident depression: a prospective cohort study of 280,032 participants. BMC Med. 2024 Apr 15;22(1):160. doi: 10.1186/s12916-024-03382-3. PMID: 38616272; PMCID: PMC11017623.
23. Shi B, Choirat C, Coull BA, VanderWeele TJ, Valeri L. CMAverse: A Suite of

Functions for Reproducible Causal Mediation Analyses. Epidemiology. 2021 Sep 1;32(5):e20-e22. doi: 10.1097/EDE.0000000000001378. PMID: 34028370.
24. Ma Y, Li D, Cui F, Wang J, Tang L, Yang Y, Liu R, Tian Y. Air pollutants, genetic susceptibility, and abdominal aortic aneurysm risk: a prospective study. Eur Heart J. 2024 Mar 27;45(12):1030-1039. doi: 10.1093/eurheartj/ehad886. PMID: 38241289.
25. Fadel M, Evanoff BA, Andersen JH, d'Errico A, Dale AM, Leclerc A, Descatha A. Not just a research method: If used with caution, can job-exposure matrices be a useful tool in the practice of occupational medicine and public health? Scand J Work Environ Health. 2020 Sep 1;46(5):552-553. doi: 10.5271/sjweh.3900. Epub 2020 May 5. PMID: 32367143; PMCID: PMC7737790.
26. Institute of Medicine (US) Committee to Assess Training Needs for Occupational Safety and Health Personnel in the United States. Safe Work in the 21st Century: Education and Training Needs for the Next Decade's Occupational Safety and Health Personnel. Washington (DC): National Academies Press (US); 2000. 2, Occupational Safety and Health Professionals. Available from: https://www.ncbi.nlm.nih.gov/books/NBK225528/.
27. Fact Sheets on the European Union European Parliament. Health and safety at work. (accessed on 22 February 2025). Available online: https://www.europarl.europa.eu/factsheets/en/sheet/56/health-and-safety-at-work.
28. Teixeira LR, Azevedo TM, Bortkiewicz A, Corrêa da Silva DT, de Abreu W, de Almeida MS, de Araujo MAN, Gadzicka E, Ivanov ID, Leppink N, Macedo MRV, de S Maciel EMG, Pawlaczyk-Łuszczyńska M, Pega F, Prüss-Üstün AM, Siedlecka J, Stevens GA, Ujita Y, Braga JU. WHO/ILO work-related burden of disease and injury: Protocol for systematic reviews of exposure to occupational noise and of the effect of exposure to occupational noise on cardiovascular disease. Environ Int. 2019 Apr;125:567-578. doi: 10.1016/j.envint.2018.09.040. Epub 2019 Jan 23. PMID: 30683322.
29. Barlow CA, Sahmel J, Paustenbach DJ, Henshaw JL. History of knowledge and evolution of occupational health and regulatory aspects of asbestos exposure science: 1900-1975. Crit Rev Toxicol. 2017 Apr;47(4):286-316. doi: 10.1080/10408444.2016.1258391. Epub 2017 Mar 22. PMID: 28328293.
30. Eng A, 't Mannetje A, McLean D, Ellison-Loschmann L, Cheng S, Pearce N. Gender differences in occupational exposure patterns. Occup Environ Med. 2011 Dec;68(12):888-94. doi: 10.1136/oem.2010.064097. Epub 2011 Apr 12. PMID: 21486991.
31. Gershon AS, Dolmage TE, Stephenson A, Jackson B. Chronic obstructive pulmonary disease and socioeconomic status: a systematic review. COPD. 2012 Jun;9(3):216-26. doi: 10.3109/15412555.2011.648030. Epub 2012 Apr 12. PMID: 22497534.
32. Ambrosino N, Bertella E. Lifestyle interventions in prevention and comprehensive management of COPD. Breathe (Sheff). 2018 Sep;14(3):186-194. doi:

10.1183/20734735.018618. PMID: 30186516; PMCID: PMC6118879.
33. Kimenai DM, Pirondini L, Gregson J, Prieto D, Pocock SJ, Perel P, Hamilton T, Welsh P, Campbell A, Porteous DJ, Hayward C, Sattar N, Mills NL, Shah ASV. Socioeconomic Deprivation: An Important, Largely Unrecognized Risk Factor in Primary Prevention of Cardiovascular Disease. Circulation. 2022 Jul 19;146(3):240-248. doi: 10.1161/CIRCULATIONAHA.122.060042. Epub 2022 Jun 24. PMID: 35748241; PMCID: PMC9287096.
34. Arnett DK, Khera A, Blumenthal RS. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: Part 1, Lifestyle and Behavioral Factors. JAMA Cardiol. 2019 Oct 1;4(10):1043-1044. doi: 10.1001/jamacardio.2019.2604. PMID: 31365022.
35. Li L, Pang S, Starnecker F, Mueller-Myhsok B, Schunkert H. Integration of a polygenic score into guideline-recommended prediction of cardiovascular disease. Eur Heart J. 2024 May 27;45(20):1843-1852. doi: 10.1093/eurheartj/ehae048. PMID: 38551411; PMCID: PMC11129792.
36. Zheng F, Gonçalves FM, Abiko Y, Li H, Kumagai Y, Aschner M. Redox toxicology of environmental chemicals causing oxidative stress. Redox Biol. 2020 Jul;34:101475. doi: 10.1016/j.redox.2020.101475. Epub 2020 Apr 18. PMID: 32336668; PMCID: PMC7327986.
37. Tkacova R. Systemic inflammation in chronic obstructive pulmonary disease: may adipose tissue play a role? Review of the literature and future perspectives. Mediators Inflamm. 2010;2010:585989. doi: 10.1155/2010/585989. Epub 2010 Apr 20. PMID: 20414465; PMCID: PMC2857618.
38. Barron HV, Cannon CP, Murphy SA, Braunwald E, Gibson CM. Association between white blood cell count, epicardial blood flow, myocardial perfusion, and clinical outcomes in the setting of acute myocardial infarction: a thrombolysis in myocardial infarction 10 substudy. Circulation. 2000 Nov 7;102(19):2329-34. doi: 10.1161/01.cir.102.19.2329. PMID: 11067784.
39. Lagrand WK, Visser CA, Hermens WT, Niessen HW, Verheugt FW, Wolbink GJ, Hack CE. C-reactive protein as a cardiovascular risk factor: more than an epiphenomenon? Circulation. 1999 Jul 6;100(1):96-102. doi: 10.1161/01.cir.100.1.96. PMID: 10393687.
40. Cui J, Chen Y, Wang HY, Wang RF. Mechanisms and pathways of innate immune activation and regulation in health and cancer. Hum Vaccin Immunother. 2014;10(11):3270-85. doi: 10.4161/21645515.2014.979640. PMID: 25625930; PMCID: PMC4514086.
41. Kusiak A, Brady G. Bifurcation of signalling in human innate immune pathways to NF-kB and IRF family activation. Biochem Pharmacol. 2022 Nov;205:115246. doi:

10.1016/j.bcp.2022.115246. Epub 2022 Sep 8. PMID: 36088989.
42. Yin Q, Fu TM, Li J, Wu H. Structural biology of innate immunity. Annu Rev Immunol. 2015;33:393-416. doi: 10.1146/annurev-immunol-032414-112258. Epub 2015 Jan 22. PMID: 25622194; PMCID: PMC4394028.
43. Ganter MT, Roux J, Miyazawa B, Howard M, Frank JA, Su G, Sheppard D, Violette SM, Weinreb PH, Horan GS, Matthay MA, Pittet JF. Interleukin-1beta causes acute lung injury via alphavbeta5 and alphavbeta6 integrin-dependent mechanisms. Circ Res. 2008 Apr 11;102(7):804-12. doi: 10.1161/CIRCRESAHA.107.161067. Epub 2008 Feb 14. PMID: 18276918; PMCID: PMC2739091.
44. Libby P. Interleukin-1 Beta as a Target for Atherosclerosis Therapy: Biological Basis of CANTOS and Beyond. J Am Coll Cardiol. 2017 Oct 31;70(18):2278-2289. doi: 10.1016/j.jacc.2017.09.028. PMID: 29073957; PMCID: PMC5687846.
45. Hanna RN, Hedrick CC. Stressing out stem cells: linking stress and hematopoiesis in cardiovascular disease. Nat Med. 2014 Jul;20(7):707-8. doi: 10.1038/nm.3631. PMID: 24999940; PMCID: PMC4467964.
46. Blanco I, Tura-Ceide O, Peinado VI, Barberà JA. Updated Perspectives on Pulmonary Hypertension in COPD. Int J Chron Obstruct Pulmon Dis. 2020 Jun 9;15:1315-1324. doi: 10.2147/COPD.S211841. PMID: 32606641; PMCID: PMC7293405.
47. Lavigne-Robichaud M, Trudel X, Talbot D, Milot A, Gilbert-Ouimet M, Vézina M, Laurin D, Dionne CE, Pearce N, Dagenais GR, Brisson C. Psychosocial Stressors at Work and Coronary Heart Disease Risk in Men and Women: 18-Year Prospective Cohort Study of Combined Exposures. Circ Cardiovasc Qual Outcomes. 2023 Oct;16(10):e009700. doi: 10.1161/CIRCOUTCOMES.122.009700. Epub 2023 Sep 19. PMID: 37724474; PMCID: PMC10573112.
48. Jurek AM, Greenland S, Maldonado G. How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? Int J Epidemiol. 2008 Apr;37(2):382-5. doi: 10.1093/ije/dym291. Epub 2008 Jan 9. PMID: 18184671.
49. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. Am J Epidemiol. 2017 Nov 1;186(9):1026-1034. doi: 10.1093/aje/kwx246. PMID: 28641372; PMCID: PMC5860371.

977

978 **Table 1. Baseline characteristics of the participants**

	COPD			CVD		
	No (N=100750)	Yes (N=2426)	Overall (N=103176)	No (N=69384)	Yes (N=18035)	Overall (N=87419)
Age at recruitment, years, mean (SD)	57.4 (6.5)	60.8 (5.7)	57.5 (6.5)	56.6 (6.4)	59.4 (6.2)	57.2 (6.5)
Sex, n (%)						
female	55903 (55.5%)	1113 (45.9%)	57016 (55.3%)	40809 (58.8%)	8220 (45.6%)	49029 (56.1%)
male	44847 (44.5%)	1313 (54.1%)	46160 (44.7%)	28575 (41.2%)	9815 (54.4%)	38390 (43.9%)
Ethnicity, n (%)						
white	98647 (97.9%)	2385 (98.3%)	101032 (97.9%)	67840 (97.8%)	17703 (98.2%)	85543 (97.9%)
other	2103 (2.1%)	41 (1.7%)	2144 (2.1%)	1544 (2.2%)	332 (1.8%)	1876 (2.1%)
Body mass index, kg/m^2, n (%)						
<25.0	38747 (38.5%)	770 (31.7%)	39517 (38.3%)	28523 (41.1%)	5849 (32.4%)	34372 (39.3%)
25.0-29.9	42388 (42.1%)	971 (40.0%)	43359 (42.0%)	28668 (41.3%)	7902 (43.8%)	36570 (41.8%)
>=30.0	19615 (19.5%)	685 (28.2%)	20300 (19.7%)	12193 (17.6%)	4284 (23.8%)	16477 (18.8%)
Polygenic risk score, mean (SD)	6.64 (0.46)	6.69 (0.46)	6.64 (0.46)	-0.19 (0.97)	-0.09 (0.97)	-0.17 (0.97)
Education, n (%)						
college/university	47514 (47.2%)	791 (32.6%)	48305 (46.8%)	33786 (48.7%)	7872 (43.6%)	41658 (47.7%)
professional qualification	27189 (27.0%)	738 (30.4%)	27927 (27.1%)	18127 (26.1%)	5206 (28.9%)	23333 (26.7%)
secondary school	19672 (19.5%)	518 (21.4%)	20190 (19.6%)	13581 (19.6%)	3484 (19.3%)	17065 (19.5%)
primary school	6375 (6.3%)	379 (15.6%)	6754 (6.5%)	3890 (5.6%)	1473 (8.2%)	5363 (6.1%)
Household income, £/year, n (%)						
<18000	13223 (13.1%)	614 (25.3%)	13837 (13.4%)	8241 (11.9%)	2913 (16.2%)	11154 (12.8%)
18000 to 30999	24382 (24.2%)	762 (31.4%)	25144 (24.4%)	16025 (23.1%)	4779 (26.5%)	20804 (23.8%)
31000 to 51999	29455 (29.2%)	617 (25.4%)	30072 (29.1%)	20418 (29.4%)	5179 (28.7%)	25597 (29.3%)
52000 to 100000	25886 (25.7%)	351 (14.5%)	26237 (25.4%)	18851 (27.2%)	4048 (22.4%)	22899 (26.2%)
>100000	7804 (7.7%)	82 (3.4%)	7886 (7.6%)	5849 (8.4%)	1116 (6.2%)	6965 (8.0%)
Townsend deprivation index, mean (SD)	-1.92 (2.70)	-1.26 (3.04)	-1.90 (2.71)	-1.90 (2.70)	-1.88 (2.74)	-1.89 (2.71)
Healthy diet score, mean (SD)	3.11 (1.36)	2.91 (1.42)	3.10 (1.36)	3.09 (1.35)	3.09 (1.38)	3.09 (1.36)
Smoking status, n (%)						
never	59360 (58.9%)	654 (27.0%)	60014 (58.2%)	41906 (60.4%)	9566 (53.0%)	51472 (58.9%)
previous	35582 (35.3%)	1228 (50.6%)	36810 (35.7%)	23375 (33.7%)	7099 (39.4%)	30474 (34.9%)
current	5808 (5.8%)	544 (22.4%)	6352 (6.2%)	4103 (5.9%)	1370 (7.6%)	5473 (6.3%)
Alcohol drinking frequency, n (%)						
>= 3 times/week	51703 (51.3%)	1215 (50.1%)	52918 (51.3%)	35616 (51.3%)	9440 (52.3%)	45056 (51.5%)
< 3 times/week	43779 (43.5%)	1045 (43.1%)	44824 (43.4%)	30311 (43.7%)	7590 (42.1%)	37901 (43.4%)
never	5268 (5.2%)	166 (6.8%)	5434 (5.3%)	3457 (5.0%)	1005 (5.6%)	4462 (5.1%)
Sleep duration, n (%)						
7-8 h/day	72991 (72.4%)	1587 (65.4%)	74578 (72.3%)	50793 (73.2%)	12723 (70.5%)	63516 (72.7%)
< 7h/day	21566 (21.4%)	632 (26.1%)	22198 (21.5%)	14632 (21.1%)	4060 (22.5%)	18692 (21.4%)
> 8h/day	6193 (6.1%)	207 (8.5%)	6400 (6.2%)	3959 (5.7%)	1252 (6.9%)	5211 (6.0%)
Physical activity, n (%)						

low	18322 (18.2%)	502 (20.7%)	18824 (18.2%)	12495 (18.0%)	3479 (19.3%)	15974 (18.3%)
moderate	43978 (43.7%)	1061 (43.7%)	45039 (43.7%)	30491 (43.9%)	7681 (42.6%)	38172 (43.7%)
high	38450 (38.2%)	863 (35.6%)	39313 (38.1%)	26398 (38.0%)	6875 (38.1%)	33273 (38.1%)

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; SD: standard deviation.

Healthy diet score was calculated based on self-reported servings of fruits, vegetables, whole grains, vegetable oil, fish, dairy, refined grains, unprocessed meats, processed meats and sugar-sweetened beverages. More details can be found in Additional file 1: Table S2.

Townsend deprivation index (including measures of unemployment, non-car ownership, non-home ownership and household overcrowding), derived from respondents' postcode was used as an indicator of area- level SES.

Education is categorized as college or University degree, secondary school (includes A levels/AS levels or equivalent, O levels/GCSEs or equivalent, CSEs or equivalent), primary school, and professional qualification (NVQ or HND or HNC or equivalent, other professional qualifications).

Figure 1. Characteristics of workplace hazard exposure

COPD

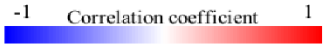
Workplace hazard	Zero proportion	T1	Mean	Median	T2	Max	chemical fumes	smoke	asbestos	paints	pesticides	diesel exhaust	noisy	cold	hot
dusty	60.8	8320	23248	15210	26312	403520	0.45*	0.28*	0.33*	0.38*	0.17*	0.41*	0.50*	0.52*	0.46*
chemical fumes	73.3	6890	21370	13260	23465	403520		0.23*	0.36*	0.46*	0.21*	0.42*	0.37*	0.35*	0.33*
smoke	42.5	8320	21890	14560	24180	308880			0.22*	0.19*	0.07*	0.30*	0.27*	0.29*	0.26*
asbestos	89.8	9360	24159	16640	29250	343200				0.36*	0.16*	0.33*	0.26*	0.26*	0.22*
paints	82.4	7800	22296	14430	25740	403520					0.20*	0.33*	0.30*	0.29*	0.25*
pesticides	95.9	5200	17576	10296	17940	223600						0.21*	0.10*	0.15*	0.12*
diesel exhaust	83.8	8025	24775	15600	29120	346632							0.34*	0.39*	0.32*
noisy	36.4	12220	26375	18720	31200	297440								0.52*	0.54*
cold	50.5	10920	23552	17030	27690	349440									0.62*
hot	39.8	12853	24864	18720	30420	291720									

Correlation analysis

CVD

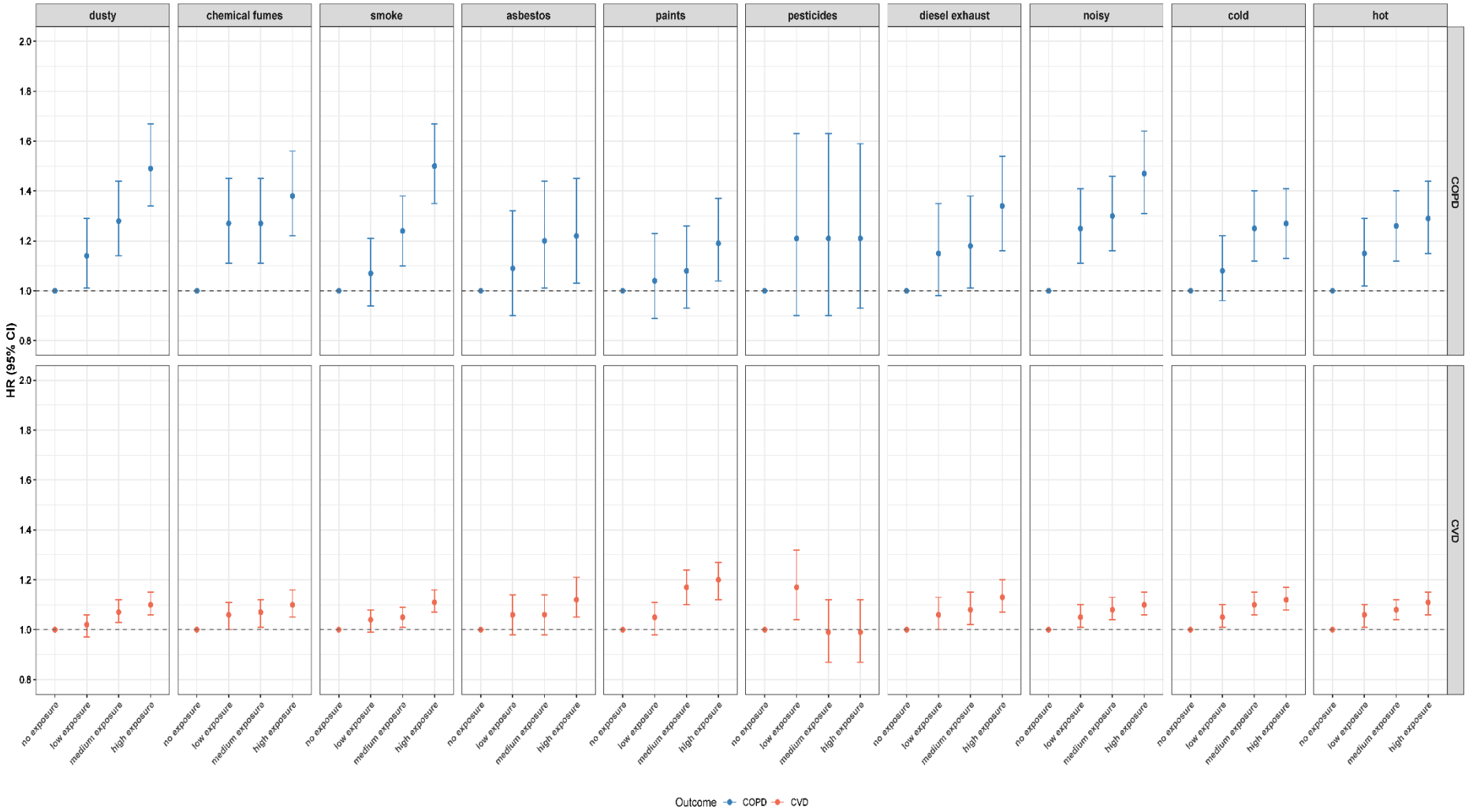
Workplace hazard	Zero proportion	T1	Mean	Median	T2	Max	chemical fumes	smoke	asbestos	paints	pesticides	diesel exhaust	noisy	cold	hot
dusty	61.3	8320	22872	14820	25740	274560	0.45*	0.28*	0.33*	0.38*	0.17*	0.41*	0.49*	0.52*	0.46*
chemical fumes	73.8	6760	20899	12870	22880	255840		0.23*	0.35*	0.45*	0.21*	0.42*	0.36*	0.35*	0.33*
smoke	42.9	8320	21432	14430	23530	362700			0.21*	0.19*	0.07*	0.29*	0.27*	0.28*	0.26*
asbestos	90.2	9360	23809	16354	28704	343200				0.36*	0.16*	0.32*	0.25*	0.26*	0.22*
paints	82.8	7280	21954	14040	24960	287404					0.19*	0.33*	0.30*	0.29*	0.25*
pesticides	96.0	5200	17388	10140	17680	203060						0.20*	0.10*	0.15*	0.12*
diesel exhaust	84.3	7800	24410	15210	28860	346632							0.33*	0.39*	0.31*
noisy	36.7	12090	25995	18720	30680	304590								0.52*	0.54*
cold	50.8	10920	23234	16640	27248	349440									0.62*
hot	40.0	12506	24595	18616	30160	291720									

Correlation analysis



Zero proportion represents the percentage of participants reporting no exposure to each hazard. T1, Mean, Median, T2, and Max are calculated among participants with non-zero exposure and reflect the cumulative LOE scores, which incorporate exposure intensity, job duration, and working hours. Correlation coefficients were estimated using Spearman's rank correlation. The heatmap color scale ranges from −1 (blue) to +1 (red). * indicates $P < 0.05$. COPD: chronic obstructive pulmonary disease; CVD: cardiovascular disease.

Figure 2. Associations between lifetime exposure to specific occupational hazard and the risks of COPD and CVD



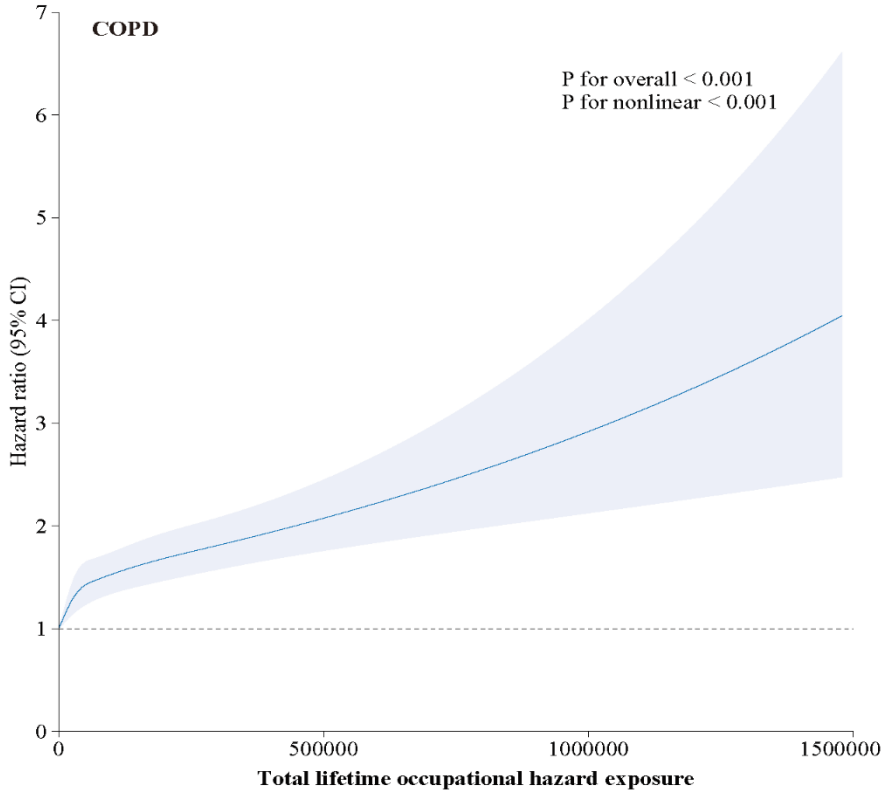
COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval.
HRs and 95% CIs for COPD and CVD across exposure levels for each occupational hazard. HRs represent category-specific estimates compared with the ‘no exposure’ group. Exposure levels (‘low,’ ‘medium,’ and ‘high’) are defined by tertiles among participants with non-zero exposure.
Models were adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.

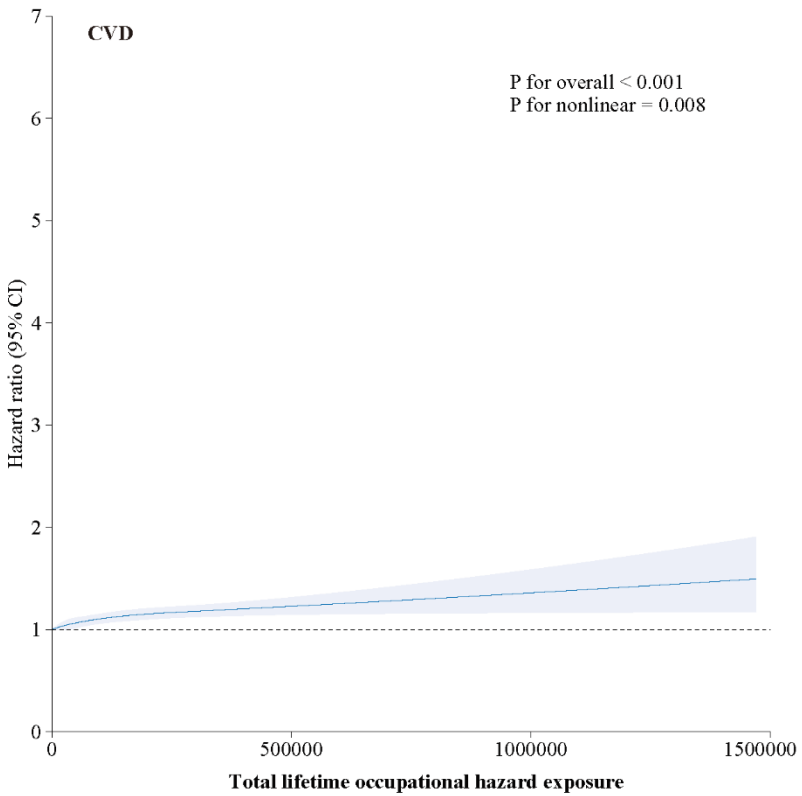
Figure 3. Associations of total lifetime occupational hazard exposure and the risks of COPD and CVD

A.

Total lifetime occupational hazard exposure	Cases	Total number	Person year	HR (95%CI)	P for trend
COPD					<0.001
no exposure	143	11528	143952	Ref.	
low exposure	525	30550	380101	1.30 (1.08–1.56)	
medium exposure	678	30549	379228	1.46 (1.22–1.76)	
high exposure	1080	30549	376179	1.83 (1.53–2.19)	
HR per exposure level increase				1.21 (1.15–1.26)	
CVD					<0.001
no exposure	1833	9970	114273	Ref.	
low exposure	4708	25817	296012	1.03 (0.97–1.09)	
medium exposure	5144	25816	293278	1.06 (1.01–1.12)	
high exposure	6350	25816	284711	1.14 (1.08–1.20)	
HR per exposure level increase				1.05 (1.03–1.06)	

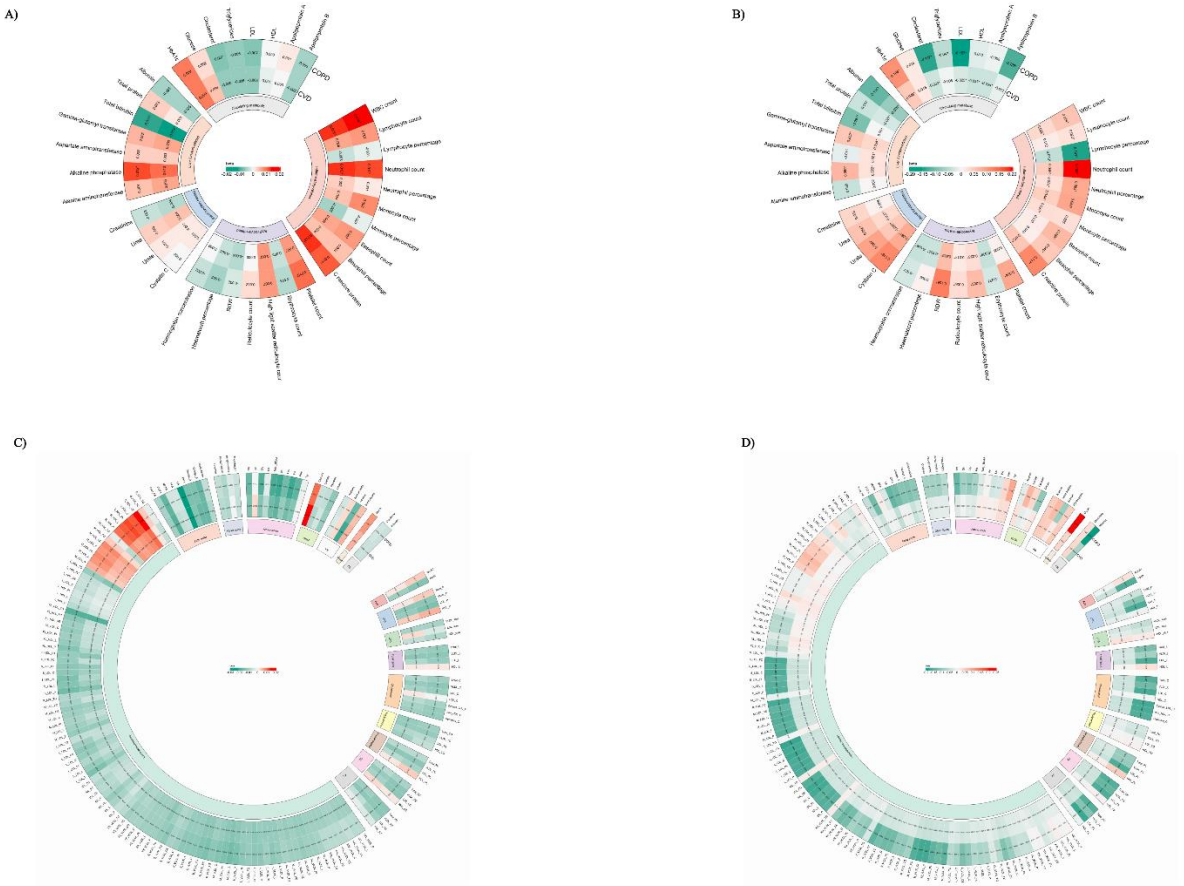
B.


COPD
P for overall < 0.001
P for nonlinear < 0.001
Hazard ratio (95% CI)
0
1
2
3
4
5
6
7
0
500000
1000000
1500000
Total lifetime occupational hazard exposure


CVD
P for overall < 0.001
P for nonlinear = 0.008
Hazard ratio (95% CI)
0
1
2
3
4
5
6
7
0
500000
1000000
1500000
Total lifetime occupational hazard exposure

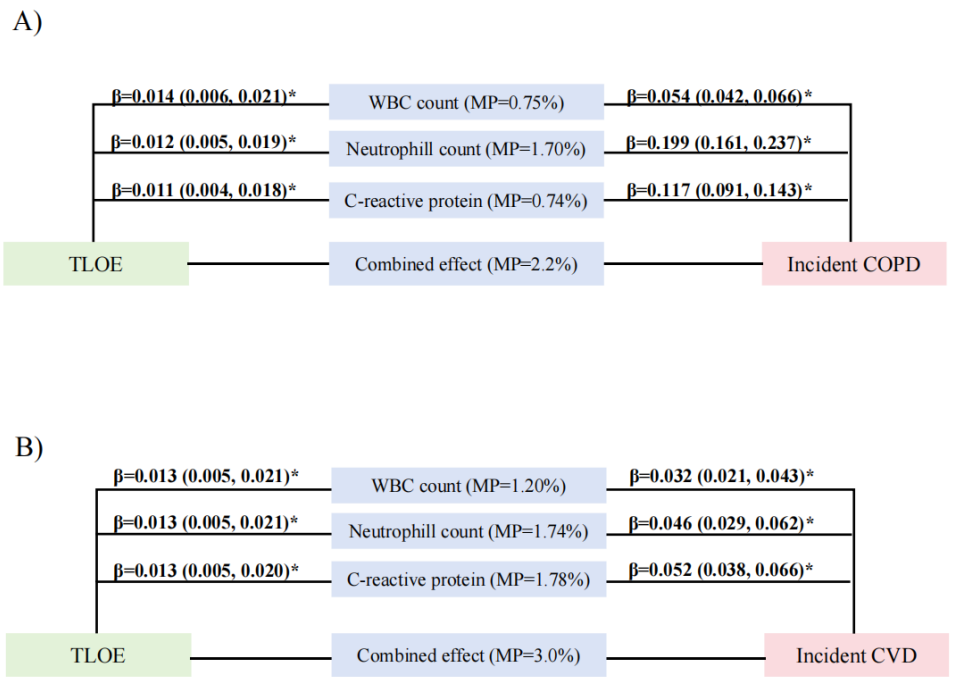
COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratio; CI, confidence interval; Ref., reference.
Panel A displays adjusted HRs and 95% CIs for COPD and CVD across four total lifetime occupational hazard exposure categories (‘no,’ ‘low,’ ‘medium,’ ‘high’), where exposure levels are defined by tertiles among participants with non-zero exposure. HRs represent category-specific estimates compared with the ‘no exposure’ group, and HRs per exposure-level increase are also reported.
Panel B presents dose–response associations estimated using restricted cubic spline Cox models with four knots placed at the 5th, 35th, 65th, and 95th percentiles of TLOE (reference is the 5th percentile). Solid lines represent adjusted HRs and shaded areas represent 95% CIs.
All models were adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.

Figure 4. Associations of biomarkers and metabolites with total lifetime occupational hazard exposure and the incidence of COPD and CVD



A) Associations of biomarkers with total lifetime occupational hazard exposure; B) Associations of biomarkers with incidence of COPD and CVD.
C) Associations of metabolites with total lifetime occupational hazard exposure; D) Associations of metabolites with incidence of COPD and CVD.
COPD: chronic obstructive pulmonary disease; CVD: Cardiovascular disease; WBC: white blood cell; RDW: red blood cell distribution width; HbA1c: glycated haemoglobin; HDL: high-density lipoprotein; VLDL: Very-low-density lipoprotein; LDL: low-density lipoprotein; Apo: Apolipoproteins; LPC: Lipoprotein particle concentrations; LPS: Lipoprotein particle sizes; CE: Cholesteryl esters; FC: Free cholesterol; GRM: Glycolysis related metabolites; KB: Ketone bodies; Inflam: Inflammation; FB: Fluid balance; CMB: Circulating metabolic biomarkers.
All biomarkers and metabolites were standardized (z-score) before entering formal analysis and total lifetime occupational hazard exposure was analyzed per exposure level increase.
Models were adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.

Figure 5. The mediation effect of biomarkers on the association between total lifetime occupational hazard exposure and incident COPD and CVD



A) The mediation effect of biomarkers on the association between total lifetime occupational hazard exposure and incident COPD; B) The mediation effect of biomarkers on the association between total lifetime occupational hazard exposure and incident CVD.

COPD: chronic obstructive pulmonary disease; CVD: Cardiovascular disease; TLOE: total lifetime occupational hazard exposure; MP: Mediated Proportion; WBC: White blood cell.

All biomarkers and metabolites were standardized (z-score) before entering formal analysis and total lifetime occupational hazard exposure was analyzed per exposure level increase.

Models were adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.

The mediators that meet the criteria: 1) independently associated with the exposure, 2) independently associated with the outcome, 3) had significant mediation effect between the exposure the outcome, 4) the sign of the nature indirect effect was the same as the sign of the total effect, 5) the relationship with outcome has clinical implications, were reported.

Table 2. Associations of total lifetime occupational hazard exposure with incident COPD, CVD, either disease, and their comorbidity (remove all COPD and CVD cases at baseline, n= 85287)

Total lifetime occupational hazard exposure	COPD	CVD	Any	Comorbidity
	HR (95%CI)	HR (95%CI)	HR (95%CI)	HR (95%CI)
no exposure	Ref.	Ref.	Ref.	Ref.
low exposure	1.14 (0.92-1.41)	1.02 (0.97-1.08)	1.03 (0.97-1.09)	1.04 (0.78-1.39)
medium exposure	1.40 (1.14-1.72)	1.06 (1.01-1.12)	1.08 (1.02-1.14)	1.16 (0.87-1.53)
high exposure	1.68 (1.37-2.05)	1.13 (1.07-1.19)	1.16 (1.10-1.22)	1.42 (1.08-1.88)
HR per exposure level increase	1.20 (1.14-1.27)	1.04 (1.03-1.06)	1.05 (1.04-1.07)	1.15 (1.06-1.24)
P for trend	<0.001	<0.001	<0.001	<0.001

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference.
"Any" refers to the incidence of either COPD or CVD during follow-up.
"Comorbidity" refers to the incidence of both COPD and CVD during follow-up.
HRs and 95% CIs are shown across four total lifetime occupational hazard exposure categories ("no," "low," "medium," "high"), with exposure levels defined by tertiles among participants with non-zero exposure. HRs represent category-specific estimates compared with the "no exposure" group, and HRs per exposure-level increase are also reported.
Models were adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS, genotyping array, and the first 10 principal components of ancestry.

Association and biological pathways between lifetime occupational exposure to workplace hazards and incident chronic obstructive pulmonary disease and cardiovascular disease in middle-aged and older adults

Additional file 1

Supplement Method. Assessment of lifetime occupational hazard exposure score

Participants were required to input their complete work history, including jobs and gap periods, from the end of their education to the present. This information was entered chronologically, allowing for a comprehensive lifetime occupational history. Once all jobs and gaps were recorded, participants reviewed their entries to ensure completeness and accuracy.

For each recorded job, participants provided the corresponding start and end years. Additionally, for each job, participants were asked to think about the place where they worked by answering ten questions: "(1) Was it very noisy? (2) Was it very cold? (3) Was it very hot? (4) Was it very dusty? (5) Was it full of chemical or other fumes? (6) Was there a lot of cigarette smoke from other people smoking? (7) Did you work with materials which contained asbestos? (8) Did you work with paints, thinners, or glues? (9) Did you work with pesticides? (10) Was there a lot of diesel exhaust?" For each question, participants could choose one of the proposed answers: "Often," "Sometimes," "Rarely/Never," or "Do not know." Responses marked as "Do not know" were treated as missing values.

To quantify participants' lifetime occupational exposure, we developed an individualized Lifetime Occupational Exposure (LOE) score. Below, we describe the calculation steps in detail.

Calculation process

Step 1: Workplace exposure scoring

1. Assign values to each workplace hazard exposure:
    - "Often": 1 point
    - "Sometimes": 0.5 point
    - "Never": 0 points
2. Calculate all workplace hazards exposure:

    Total workplace hazards exposure = Sum of the scores for all 10 hazards (possible range: 0-10)

Step 2: Calculation of Length Worked

1. Calculate the number of years worked for each job:
    - Years Worked = Year job ended - Year job started + 1
    - If a job started and ended in the same year, set the years worked as 1 year to account for at least partial work during the year.
2. Determine weekly working hours:
    - 15 to less than 20 hours: 15 hours

- 20 to less than 30 hours: 20 hours
- 30 to 40 hours: 30 hours
- Over 40 hours: 40 hours

3. Calculate the length worked for each job

   Length worked for each job (hours) = Years Worked * Weekly Hours * 52

Step 4: Calculation of Cumulative Exposure Score for each job

Each hazard: Cumulative Exposure Score = each hazard exposure score × Years Worked × length worked for each job

Total hazards: Cumulative Exposure Score = total hazards exposure score × Years Worked × length worked for each job

Step 5: Calculation of total LOE score

LOE Score = Sum of the cumulative exposure scores for all jobs
(both for each hazard and total hazards)

Sensitivity analysis

In this study, we conducted a sensitivity analysis focused on the calculation of "Years Worked" for each job. We modified the years worked calculation from "Year job ended - Year job started + 1" to "Year job ended - Year job started + 0.5". This adjustment provides a more conservative estimate for partial-year employment periods, as most individuals begin or end jobs at various points throughout a calendar year rather than precisely on January 1st or December 31st.

Table S1. Characteristics of 82 SNPs associated with genetic predisposition to COPD

rsID	Marker name	Locus	Closest gene	Risk/alternative allele	OR	P value
rs9435731	1:17306029:C:A	1p36.13	MFAP2	A/C	1.06	6.60E-10
rs76841360	1:40060025:G:A	1p34.3	PABPC4	A/G	1.08	5.00E-10
rs4660861	1:45946636:G:T	1p34.1	TESK2	G/T	1.06	4.40E-08
rs72673419	1:60913143:C:T	1p32.1	C1orf87	T/C	1.14	4.00E-09
rs629619	1:111738108:C:T	1p13.3	DENND2D	T/C	1.08	2.90E-10
rs3009947	1:218689155:T:C	1q41	TGFB2	C/T	1.06	4.00E-09
rs11118406	1:219924894:T:A	1q41	SLC30A10	T/A	1.08	4.10E-11
rs11579382	1:239901006:G:C	1q43	CHRM3	C/G	1.06	6.50E-09
rs955277	2:9290357:C:T	2p25.1	ASAP2	T/C	1.07	1.90E-10
rs10929386	2:15906179:C:T	2p24.3	DDX1	C/T	1.06	9.10E-09
rs12466981	2:42433247:C:T	2p21	EML4	C/T	1.06	4.90E-08
rs72902175	2:157013035:C:T	2q24.1	NR4A2	T/C	1.09	1.60E-08
rs2571445	2:218683154:A:G	2q35	TNS1	A/G	1.07	3.40E-12
rs16825267	2:229569919:C:G	2q36.3	PID1	C/G	1.19	1.80E-20
rs62191105	2:239872704:C:T	2q37.3	TWIST2	C/T	1.09	2.60E-12
rs2442776	3:11640601:G:A	3p25.3	VGLL4	G/A	1.09	2.00E-10
rs1529672	3:25520582:C:A	3p24.2	RARB	C/A	1.09	2.50E-11
rs13073544	3:29472412:G:C	3p24.1	RBMS3	C/G	1.06	2.00E-08
rs17759204	3:55158224:A:G	3p14.3	CACNA2D3	G/A	1.07	8.80E-11
rs62259026	3:57746515:C:T	3p14.3	SLMAP	C/T	1.07	2.40E-08
rs4093840	3:123077042:T:A	3q21.1	ADCY5	A/T	1.06	3.90E-10
rs2955083	3:127961178:T:A	3q21.3	EEFSEC	A/T	1.13	3.50E-15
rs7650602	3:141147414:T:C	3q23	ZBTB38	C/T	1.06	4.90E-08
rs7642001	3:168746145:G:A	3q26.2	MECOM	A/G	1.08	1.10E-14
rs4585380	4:75673363:G:A	4q13.3	BTC	G/A	1.07	3.40E-09
rs7671261	4:89883818:G:A	4q22.1	FAM13A	A/G	1.09	1.40E-18
rs34712979	4:106819053:G:A	4q24	NPNT	A/G	1.18	3.00E-46
rs13140176	4:145489098:A:G	4q31.21	HHIP	A/G	1.18	4.10E-59
rs1551943	5:52195033:G:A	5q11.2	ITGA1	A/G	1.08	5.80E-10
rs34651	5:72144005:C:T	5q13.2	TNPO1	C/T	1.11	3.00E-08
rs153916	5:95036700:C:T	5q15	SPATA9	T/C	1.06	6.30E-10
rs62375246	5:132439010:T:A	5q31.1	HSPA4	A/T	1.06	2.20E-08
rs10037493	5:147854970:C:T	5q32	HTR4	C/T	1.13	2.60E-33
rs979453	5:150595073:A:G	5q33.1	CCDC69	G/A	1.06	1.40E-08
rs10866659	5:156937043:A:G	5q33.3	ADAM19	G/A	1.09	1.20E-16
rs12519165	5:170901586:A:T	5q35.1	FGF18	A/T	1.07	1.10E-09
rs1334576	6:7211818:G:A	6p24.3	RREB1	A/G	1.06	1.20E-08
rs9350191	6:19842661:C:T	6p22.3	ID4	T/C	1.12	5.10E-14
rs13198656	6:22004909:C:T	6p22.3	PRL	T/C	1.06	1.20E-09
rs3095329	6:30693816:A:G	6p21.33	IER3	G/A	1.12	2.10E-21

rs2070600	6:32151443:C:T	6p21.32	AGER	C/T	1.21	1.10E-17
rs2806356	6:109266255:T:C	6q21	ARMC2	C/T	1.1	2.90E-15
rs1631199	6:117259673:C:G	6q22.1	RFX6	G/C	1.06	7.60E-09
rs646695	6:140280398:T:C	6q24.1	CITED2	C/T	1.08	4.60E-11
rs9399401	6:142668901:T:C	6q24.1	ADGRG6	T/C	1.16	1.60E-40
rs798565	7:2752152:G:A	7p22.3	AMZ1	G/A	1.07	3.90E-09
rs2040732	7:20418134:C:T	7p21.1	ITGB8	C/T	1.06	6.90E-09
rs2897075	7:99630342:C:T	7q22.1	ZKSCAN1	C/T	1.07	7.30E-10
rs9329170	8:8697658:C:G	8p23.1	MFHAS1	C/G	1.1	3.60E-10
rs10114763	9:4143749:A:T	9p24.2	GLIS3	T/A	1.07	8.70E-13
rs156394	9:23588684:T:C	9p21.3	ELAVL2	T/C	1.07	3.80E-12
rs7866939	9:85126163:T:C	9q21.32	RASEF	C/T	1.06	1.80E-08
rs10760580	9:101661650:G:A	9q22.33	COL15A1	G/A	1.07	1.20E-10
rs803923	9:119401650:G:A	9q33.1	ASTN2	A/G	1.06	2.70E-08
rs7068966	10:12277992:C:T	10p13	CDC123	C/T	1.1	6.20E-23
rs2579762	10:78318879:A:C	10q22.3	LRMDA	C/A	1.06	2.60E-10
rs721917	10:81706324:A:G	10q22.3	SFTPD	G/A	1.06	2.20E-08
rs1570221	10:105656874:G:A	10q24.33	STN1	A/G	1.06	2.20E-08
rs4757118	11:13171236:C:T	11p15.2	ARNTL	T/C	1.06	3.80E-09
rs117261012	11:86444761:A:G	11q14.2	PRSS23	G/A	1.09	6.90E-10
rs11049386	12:28320536:T:A	12p11.22	CCDC91	T/A	1.06	2.70E-08
rs7307510	12:96237570:T:C	12q23.1	SNRPF	C/T	1.08	2.60E-09
rs7958945	12:115947901:A:G	12q24.21	MED13L	G/A	1.06	1.00E-09
rs9525927	13:44842503:G:A	13q14.11	SERP2	G/A	1.08	2.80E-09
rs72699855	14:93105953:G:C	14q32.12	RIN3	G/C	1.08	4.80E-09
rs72731149	15:49984710:G:C	15q21.2	DTWD1	G/C	1.12	8.30E-09
rs1441358	15:71612514:T:G	15q23	THSD4	G/T	1.13	7.40E-33
rs55676755	15:78898932:C:G	15q25.1	CHRNA3	G/C	1.11	2.70E-26
rs10152300	15:84392907:G:A	15q25.2	ADAMTSL3	G/A	1.08	4.20E-12
rs56134392	16:10709013:T:C	16p13.13	TEKT5	C/T	1.06	4.50E-08
rs8044657	16:58022625:G:A	16q21	TEPP	G/A	1.11	1.10E-08
rs4888379	16:75340231:A:T	16q23.1	CFDP1	T/A	1.1	5.90E-21
rs8080772	17:28413129:T:C	17q11.2	EFCAB5	T/C	1.06	1.40E-08
rs34727469	17:36835079:C:T	17q12	RPL23	T/C	1.09	1.10E-08
rs62065216	17:38218773:G:A	17q21.1	THRA	A/G	1.06	8.10E-09
rs12185268	17:43923683:A:G	17q21.31	SPPL2C	G/A	1.08	9.90E-10
rs11655567	17:69216687:T:C	17q24.3	SOX9	C/T	1.06	1.90E-09
rs647097	18:8808464:T:C	18p11.22	MTCL1	C/T	1.08	1.00E-11
rs72626215	19:46294136:G:A	19q13.32	DMWD	G/A	1.07	1.70E-09
rs2096468	21:35661745:A:C	21q22.11	KCNE2	A/C	1.06	4.00E-08
rs9617650	22:18488883:G:C	22q11.21	MICAL3	G/C	1.08	4.40E-10
rs73158393	22:33335386:C:G	22q12.3	SYN3	C/G	1.07	7.70E-09

Table S2. Definition of each component of a healthy diet score

Components	Goal (1 point)	Amount per serving	Field IDs
fruits	≥ 3 servings/day	1 piece of fresh fruit 5 pieces of dried fruit	1309, 1319
vegetables	≥ 3 servings/day	3 heaped tablespoons	1289, 1299
whole grains	≥ 3 servings/day	1 slice of whole-grain bread 1 cup of whole-grain cereal	1438, 1448, 1458, 1468
vegetable oil	≥ 2 servings/day	1 serving/day if in combination with eating at least 2 slices of bread (ID 1438)	1428 (Flora Pro-Active/Benecol spread), 2654 (Flora Pro-Active/Benecol, soft margarine, olive oil based, polyunsaturated/sunflower oil based, other low/reduced fat spread), 1438 (bread slices/week)
(shell) fish	≥ 2 servings/week	Once/week	1329, 1339
dairy	≥ 2 servings/day	1 glass/day if consumption any type of milk 1 piece of cheese	1408, 1418
refined grains	≤ 2 servings/day	1 slice of bread 1 bowl of cereal	1438, 1448 (white, brown, other bread slices/week) 1458, 1468 (biscuit, other cereals/week)
unprocessed meats	≤ 2 servings/week	once/week (including poultry, beef, lamb, and pork) 0 pieces/day if indicated having never eaten meat 1 piece/day	1359, 1369, 1379, 1389, 3680
processed meats	≤ 1 servings/week	0 pieces/day if indicated having never eaten meat	1349, 3680
sugar-sweetened beverages	don't drink	0 serving	6144

Data on food consumption were derived from the baseline questionnaire. If participants achieved the intake goal of each diet component, they were considered to have an adequate intake and get one point. The points were then accumulated to calculate the final healthy diet score. A higher score indicated a healthier diet pattern.

Table S3. Associations between total lifetime occupational hazard exposure and the risk of COPD stratified by various factors

		Low exposure	Medium exposure	High exposure	Per exposure level increase	P for trend
age	<60 years	1.44 (1.04-1.98)	1.88 (1.38-2.57)	2.71 (1.99-3.68)	1.38 (1.29-1.49)	<0.001
	≥60 years	1.41 (1.12-1.77)	1.75 (1.41-2.19)	2.69 (2.17-3.33)	1.39 (1.32-1.47)	<0.001
sex	female	1.37 (1.08-1.75)	1.68 (1.32-2.14)	2.07 (1.62-2.66)	1.25 (1.17-1.34)	<0.001
	male	1.41 (1.06-1.88)	1.85 (1.40-2.43)	3.02 (2.33-3.90)	1.47 (1.38-1.57)	<0.001
PRS	≤ median	1.20 (0.92-1.55)	1.43 (1.11-1.84)	2.42 (1.89-3.09)	1.39 (1.31-1.49)	<0.001
	> median	1.58 (1.21-2.05)	2.09 (1.62-2.71)	3.01 (2.34-3.88)	1.41 (1.33-1.49)	<0.001
education level	college/university	1.25 (0.95-1.63)	1.77 (1.36-2.29)	2.07 (1.59-2.69)	1.29 (1.20-1.38)	<0.001
	other	1.40 (1.08-1.81)	1.58 (1.23-2.04)	2.62 (2.05-3.34)	1.39 (1.32-1.47)	<0.001
household income	≥52000	1.55 (1.05-2.30)	1.94 (1.32-2.87)	2.95 (2.01-4.31)	1.40 (1.27-1.55)	<0.001
	< 52000	1.26 (1.02-1.55)	1.54 (1.26-1.89)	2.33 (1.91-2.84)	1.35 (1.29-1.42)	<0.001
TDI	≤ median	1.29 (0.99-1.68)	1.68 (1.30-2.17)	2.22 (1.72-2.85)	1.31 (1.23-1.40)	<0.001
	> median	1.45 (1.12-1.88)	1.79 (1.39-2.31)	3.08 (2.41-3.93)	1.47 (1.38-1.56)	<0.001
healthy diet score	≤ median	1.38 (1.10-1.75)	1.87 (1.49-2.34)	2.89 (2.32-3.61)	1.44 (1.37-1.52)	<0.001
	> median	1.36 (1.01-1.84)	1.55 (1.15-2.09)	2.38 (1.78-3.18)	1.33 (1.24-1.43)	<0.001
Physical activity	high	1.23 (0.89-1.68)	1.49 (1.10-2.03)	2.45 (1.82-3.29)	1.40 (1.30-1.51)	<0.001
	other	1.46 (1.16-1.83)	1.89 (1.51-2.37)	2.88 (2.32-3.58)	1.42 (1.34-1.49)	<0.001
drinking status	never	0.81 (0.45-1.48)	0.78 (0.43-1.42)	1.66 (0.95-2.90)	1.32 (1.11-1.55)	0.031
	other	1.44 (1.18-1.75)	1.85 (1.53-2.24)	2.82 (2.34-3.39)	1.41 (1.35-1.47)	<0.001
smoking status	never	1.33 (0.98-1.80)	1.62 (1.20-2.18)	2.05 (1.52-2.76)	1.26 (1.16-1.36)	<0.001
	other	1.21 (0.96-1.52)	1.46 (1.17-1.84)	2.18 (1.75-2.72)	1.33 (1.27-1.41)	<0.001
sleep duration	7-8 h/day	1.31 (1.05-1.62)	1.63 (1.32-2.01)	2.51 (2.05-3.08)	1.38 (1.30-1.45)	<0.001
	other	1.56 (1.08-2.24)	2.00 (1.41-2.84)	3.03 (2.15-4.27)	1.42 (1.32-1.54)	<0.001

COPD, chronic obstructive pulmonary disease; PRS, polygenic risk score; TDI, Townsend Deprivation Index; HR, hazard ratios; CI, confidence interval. The results are expressed as hazard ratios (HRs) with 95% confidence intervals (CIs). No-exposure was regarded as “reference”.

Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, COPD PRS, genotyping array, and the first 10 principal components of ancestry. When the analysis is stratified by a specific variable, that variable is not included in the model.

Table S4. Associations between total lifetime occupational hazard exposure and the risk of CVD stratified by various factors

		Low exposure	Medium exposure	High exposure	Per exposure level increase	P for trend
age	<60 years	1.05 (0.96-1.14)	1.13 (1.04-1.23)	1.41 (1.30-1.52)	1.14 (1.11-1.16)	<0.001
	≥60 years	0.96 (0.90-1.03)	1.05 (0.97-1.12)	1.25 (1.16-1.34)	1.10 (1.08-1.12)	<0.001
sex	female	1.04 (0.96-1.12)	1.08 (1.00-1.17)	1.15 (1.06-1.25)	1.05 (1.02-1.07)	<0.001
	male	1.01 (0.93-1.09)	1.10 (1.02-1.19)	1.26 (1.18-1.35)	1.10 (1.08-1.12)	<0.001
PRS	≤ median	1.01 (0.94-1.09)	1.11 (1.02-1.20)	1.36 (1.26-1.46)	1.13 (1.10-1.15)	<0.001
	> median	0.95 (0.88-1.03)	1.03 (0.95-1.10)	1.27 (1.19-1.37)	1.11 (1.09-1.14)	<0.001
education level	college/university	0.97 (0.90-1.04)	1.06 (0.99-1.14)	1.22 (1.13-1.31)	1.08 (1.06-1.11)	<0.001
	other	0.99 (0.91-1.07)	1.05 (0.97-1.13)	1.33 (1.23-1.44)	1.13 (1.11-1.16)	<0.001
household income	≥52000	0.98 (0.89-1.07)	1.07 (0.98-1.17)	1.26 (1.15-1.38)	1.09 (1.07-1.13)	<0.001
	< 52000	0.96 (0.90-1.03)	1.03 (0.96-1.10)	1.28 (1.20-1.36)	1.11 (1.09-1.13)	<0.001
TDI	≤ median	0.97 (0.90-1.04)	1.08 (1.00-1.16)	1.29 (1.20-1.39)	1.12 (1.09-1.14)	<0.001
	> median	1.00 (0.93-1.08)	1.06 (0.98-1.14)	1.34 (1.24-1.45)	1.13 (1.10-1.15)	<0.001
healthy diet score	≤ median	0.96 (0.90-1.03)	1.06 (0.99-1.13)	1.34 (1.26-1.44)	1.14 (1.11-1.16)	<0.001
	> median	1.01 (0.93-1.10)	1.08 (0.99-1.17)	1.27 (1.17-1.38)	1.10 (1.07-1.12)	<0.001
Physical activity	high	1.06 (0.97-1.17)	1.14 (1.04-1.25)	1.40 (1.28-1.53)	1.13 (1.11-1.16)	<0.001
	other	0.94 (0.88-1.01)	1.03 (0.96-1.10)	1.27 (1.19-1.36)	1.11 (1.09-1.13)	<0.001
drinking status	never	1.06 (0.83-1.34)	1.15 (0.91-1.46)	1.40 (1.11-1.77)	1.13 (1.06-1.21)	<0.001
	other	0.98 (0.93-1.03)	1.06 (1.00-1.12)	1.31 (1.24-1.38)	1.12 (1.10-1.14)	<0.001
smoking status	never	0.98 (0.91-1.05)	1.08 (1.01-1.15)	1.25 (1.17-1.34)	1.09 (1.07-1.12)	<0.001
	other	0.94 (0.86-1.03)	0.99 (0.91-1.08)	1.27 (1.16-1.38)	1.12 (1.10-1.15)	<0.001
sleep duration	7-8 h/day	0.97 (0.91-1.03)	1.07 (1.01-1.14)	1.32 (1.24-1.40)	1.12 (1.10-1.14)	<0.001
	other	1.02 (0.91-1.13)	1.04 (0.94-1.16)	1.29 (1.17-1.43)	1.11 (1.08-1.14)	<0.001

CVD, cardiovascular disease; PRS, polygenic risk score; TDI, Townsend Deprivation Index; HR, hazard ratios; CI, confidence interval. The results are expressed as hazard ratios (HRs) with 95% confidence intervals (CIs). No-exposure was regarded as “reference”.

Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, CVD PRS, genotyping array, and the first 10 principal components of ancestry. When the analysis is stratified by a specific variable, that variable is not included in the model.

Table S5. Associations of total lifetime occupational hazard exposure and the risks of CHD, AF, HF and stroke

Total lifetime occupational hazard exposure	Excluding all CVDs at baseline		Excluding only the index disease at baseline	
	HR (95%CI)	P for trend	HR (95%CI)	P for trend
CHD		0.008		0.002
no exposure	Ref.		Ref.	
low exposure	0.98 (0.88-1.10)		1.02 (0.93-1.12)	
medium exposure	1.03 (0.93-1.15)		1.03 (0.94-1.14)	
high exposure	1.11 (1.01-1.22)		1.12 (1.02-1.23)	
HR per exposure level increase	1.04 (1.01-1.07)		1.04 (1.02-1.07)	
AF		0.106		0.009
no exposure	Ref.		Ref.	
low exposure	0.98 (0.87-1.10)		0.98 (0.89-1.08)	
medium exposure	1.01 (0.90-1.13)		1.01 (0.92-1.11)	
high exposure	1.08 (0.98-1.19)		1.10 (1.01-1.21)	
HR per exposure level increase	1.03 (0.99-1.06)		1.04 (1.01-1.07)	
HF		0.822		0.312
no exposure	Ref.		Ref.	
low exposure	0.93 (0.76-1.14)		0.94 (0.80-1.10)	
medium exposure	0.96 (0.79-1.17)		0.95 (0.82-1.12)	
high exposure	1.01 (0.84-1.23)		1.03 (0.88-1.20)	
HR per exposure level increase	1.00 (0.94-1.06)		1.02 (0.98-1.07)	
Stroke		0.735		0.810
no exposure	Ref.		Ref.	
low exposure	0.92 (0.75-1.13)		0.97 (0.81-1.15)	
medium exposure	0.98 (0.81-1.20)		0.97 (0.83-1.17)	

high exposure	1.01 (0.83-1.23)		1.04 (0.87-1.23)	
HR per exposure level increase	1.01 (0.95-1.07)		1.00 (0.94-1.05)	

CHD, coronary heart disease; AF, atrial fibrillation; HF, heart failure; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference.
Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, CVD PRS, genotyping array, and the first 10 principal components of ancestry.

Table S6. Associations of total lifetime occupational hazard exposure and prevalent CVD at baseline with the risk of COPD

Prevalent CVD at baseline	Total lifetime occupational hazard exposure	Cases	Total number	Person year	Joint analysis	Subgroup analysis	P for interaction
					HR (95%CI)	HR (95%CI)	
no	no exposure	115	9814	122639	Ref.	Ref.	<0.001
	low exposure	374	25672	320208	1.15（0.93–1.42）	1.16（0.94–1.43）	
	medium exposure	498	25337	315135	1.38（1.13–1.69）	1.39（1.14–1.71）	
	high exposure	712	24464	302631	1.65（1.35–2.02）	1.67（1.36–2.04）	
yes	no exposure	28	1714	21313	1.09（0.72–1.65）	Ref.	
	low exposure	151	4878	59894	2.02（1.58–2.57）	1.79（1.20–2.69）	
	medium exposure	180	5212	64093	2.30（1.90–2.72）	2.01（1.34–3.02）	
	high exposure	368	6085	73549	2.71（2.19–3.37）	2.41（1.63–3.56）	

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference. Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, COPD PRS, genotyping array, and the first 10 principal components of ancestry.

Table S7. Associations of total lifetime occupational hazard exposure and prevalent COPD at baseline with the risk of CVD

Prevalent COPD at baseline	Total lifetime occupational hazard exposure	Cases	Total number	Person year	Joint analysis	Subgroup analysis	P for interaction
					HR (95%CI)	HR (95%CI)	
no	no exposure	1804	9814	112501	Ref.	Ref.	<0.001
	low exposure	4583	25292	290233	1.02（0.97–1.08）	1.02（0.97–1.08）	
	medium exposure	4979	25172	286256	1.06（1.01–1.12）	1.06（1.01–1.12）	
	high exposure	6077	25009	276367	1.13（1.07–1.20）	1.13（1.07–1.19）	
yes	no exposure	29	156	1771	1.00（0.69–1.45）	Ref.	
	low exposure	125	525	5779	1.29（1.07–1.54）	1.28（0.85–1.93）	
	medium exposure	165	644	7022	1.38（1.19–1.60）	1.32（0.88–1.98）	
	high exposure	273	807	8344	1.50（1.32–1.70）	1.64（1.10–2.44）	

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference. Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, CVD PRS, genotyping array, and the first 10 principal components of ancestry.

Table S8. Associations of total lifetime occupational hazard exposure and the risks of COPD and CVD (using 0.5-year adjustment for job duration)

Total lifetime occupational hazard exposure	Cases	Total number	Person year	HR (95%CI)	P for trend
COPD					<0.001
no exposure	143	11528	143952	Ref.	
low exposure	526	30550	380085	1.30 (1.08-1.56)	
medium exposure	677	30549	379232	1.46 (1.22-1.75)	
high exposure	1080	30549	376192	1.83 (1.53-2.19)	
HR per exposure level increase				1.21 (1.15-1.26)	
CVD					<0.001
no exposure	1833	9970	114273	Ref.	
low exposure	4707	25817	296015	1.03 (0.97-1.08)	
medium exposure	5144	25816	293274	1.06 (1.01-1.12)	
high exposure	6351	25816	284713	1.14 (1.08-1.20)	
HR per exposure level increase				1.05 (1.03-1.06)	

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference. Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.

Table S9. Associations of total lifetime occupational hazard exposure and the risks of COPD and CVD (remove cases within first 2 years, N=102945, 84969, respectively)

Total lifetime occupational hazard exposure	Cases	Total number	Person year	HR (95%CI)	P for trend
COPD					<0.001
no exposure	131	11516	143938	Ref.	
low exposure	475	30500	380053	1.28 (1.06-1.56)	
medium exposure	614	30485	379157	1.46 (1.20-1.76)	
high exposure	975	30444	376074	1.83 (1.52-2.21)	
HR per exposure level increase				1.21 (1.15-1.27)	
CVD					<0.001
no exposure	1597	9734	114027	Ref.	
low exposure	4086	25195	295375	1.03 (0.97-1.09)	
medium exposure	4441	25113	292557	1.05 (1.00-1.12)	
high exposure	5461	24927	283805	1.12 (1.06-1.19)	
HR per exposure level increase				1.04 (1.02-1.06)	

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference. Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.

Table S10. Associations of total lifetime occupational hazard exposure and the risks of COPD (remove cases within FEV1/FVC<0.70 at baseline, N=90575)

Total lifetime occupational hazard exposure	Cases	Total number	Person year	HR (95%CI)	P for trend
no exposure	93	10185	127372	Ref.	<0.001
low exposure	341	27054	337443	1.29 (1.04-1.62)	
medium exposure	420	26852	334448	1.43 (1.14-1.79)	
high exposure	658	26484	327778	1.86 (1.49-2.32)	
HR per exposure level increase				1.22 (1.15-1.29)	

COPD, chronic obstructive pulmonary disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference.

Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, COPD PRS, genotyping array, and the first 10 principal components of ancestry.

Table S11. Associations of total lifetime occupational hazard exposure and the risks of COPD (excluding incident asthma cases during follow-up, N=100980)

Total lifetime occupational hazard exposure	Cases	Total number	Person year	HR (95%CI)	P for trend
no exposure	131	11341	141636	Ref.	<0.001
low exposure	471	29948	372752	1.28 (1.05-1.55)	
medium exposure	621	29907	371427	1.47 (1.22-1.78)	
high exposure	982	29784	367130	1.84 (1.53-2.22)	
HR per exposure level increase				1.21 (1.16-1.27)	

COPD, chronic obstructive pulmonary disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference.

Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, COPD PRS, genotyping array, and the first 10 principal components of ancestry.

Table S12. Associations of total lifetime occupational hazard exposure and the risks of COPD and CVD (covariates without imputation)

Total lifetime occupational hazard exposure	Cases	Total number	Person year	HR (95%CI)	P for trend
COPD					<0.001
no exposure	143	11528	143952	Ref.	
low exposure	525	30550	380101	1.30 (1.08–1.56)	
medium exposure	678	30549	379228	1.46 (1.22–1.76)	
high exposure	1080	30549	376179	1.83 (1.53–2.19)	
HR per exposure level increase				1.21 (1.15–1.26)	
CVD					<0.001
no exposure	1833	9970	114273	Ref.	
low exposure	4708	25817	296012	1.03 (0.97–1.09)	
medium exposure	5144	25816	293278	1.06 (1.01–1.12)	
high exposure	6350	25816	284711	1.14 (1.08–1.20)	
HR per exposure level increase				1.05 (1.03–1.06)	

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference. Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.

Table S13. Associations of total lifetime occupational hazard exposure and the risks of COPD and CVD (further adjusted for prevalent diseases)

Total lifetime occupational hazard exposure	Main model	P for trend	Model 2a	P for trend
	HR (95%CI)		HR (95%CI)	
COPD		<0.001		<0.001
no exposure	Ref.		Ref.	
low exposure	1.30 (1.08–1.56)		1.24 (1.04–1.50)	
medium exposure	1.46 (1.22–1.76)		1.40 (1.18–1.70)	
high exposure	1.83 (1.53–2.19)		1.71 (1.43–2.04)	
HR per exposure level increase	1.21 (1.15–1.26)		1.18 (1.13–1.24)	
CVD		<0.001		<0.001
no exposure	Ref.		Ref.	
low exposure	1.03 (0.97–1.09)		1.03 (0.97–1.08)	
medium exposure	1.06 (1.01–1.12)		1.05 (1.00–1.11)	
high exposure	1.14 (1.08–1.20)		1.13 (1.07–1.19)	
HR per exposure level increase	1.05 (1.03–1.06)		1.04 (1.03–1.06)	

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference.
Main model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.
Model 2a was further adjusted for prevalent depression, anxiety, asthma, hypertension, diabetes, COPD (for incident CVD) and CVD (for incident COPD) at baseline.

Table S14. Associations of total lifetime occupational hazard exposure and the risks of COPD and CVD (further adjusted for medications)

Total lifetime occupational hazard exposure	Main model	P for trend	Model 2b	P for trend
	HR (95%CI)		HR (95%CI)	
COPD		<0.001		<0.001
no exposure	Ref.		Ref.	
low exposure	1.30 (1.08–1.56)		1.30 (1.08–1.56)	
medium exposure	1.46 (1.22–1.76)		1.46 (1.22–1.76)	
high exposure	1.83 (1.53–2.19)		1.83 (1.53–2.19)	
HR per exposure level increase	1.21 (1.15–1.26)		1.21 (1.15–1.26)	
CVD		<0.001		<0.001
no exposure	Ref.		Ref.	
low exposure	1.03 (0.97–1.09)		1.03 (0.97–1.09)	
medium exposure	1.06 (1.01–1.12)		1.06 (1.01–1.12)	
high exposure	1.14 (1.08–1.20)		1.14 (1.08–1.20)	
HR per exposure level increase	1.05 (1.03–1.06)		1.05 (1.03–1.06)	

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference.
Main model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.
Model 2b was further adjusted for aspirin, cholesterol-lowering drugs, blood pressure medications, and insulin.

Table S15. Associations of total lifetime occupational hazard exposure and the risks of COPD and CVD (further adjusted for family history of disease)

Total lifetime occupational hazard exposure	Main model	P for trend	Model 2c	P for trend
	HR (95%CI)		HR (95%CI)	
COPD		<0.001		<0.001
no exposure	Ref.		Ref.	
low exposure	1.30 (1.08–1.56)		1.29 (1.07–1.56)	
medium exposure	1.46 (1.22–1.76)		1.45 (1.21–1.74)	
high exposure	1.83 (1.53–2.19)		1.80 (1.51–2.16)	
HR per exposure level increase	1.21 (1.15–1.26)		1.20 (1.14–1.25)	
CVD		<0.001		<0.001
no exposure	Ref.		Ref.	
low exposure	1.03 (0.97–1.09)		1.02 (0.97–1.08)	
medium exposure	1.06 (1.01–1.12)		1.06 (1.01–1.12)	
high exposure	1.14 (1.08–1.20)		1.14 (1.08–1.19)	
HR per exposure level increase	1.05 (1.03–1.06)		1.05 (1.03–1.06)	

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference.
Main model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.
Model 2c was further adjusted for family history of diabetes, cardiovascular disease, hypertension, and lung disease.

Table S16. Associations of total lifetime occupational hazard exposure and the risks of COPD and CVD (further adjusted for environmental factors)

Total lifetime occupational hazard exposure	Main model	P for trend	Model 2d	P for trend
	HR (95%CI)		HR (95%CI)	
COPD		<0.001		<0.001
no exposure	Ref.		Ref.	
low exposure	1.30 (1.08–1.56)		1.29 (1.07–1.55)	
medium exposure	1.46 (1.22–1.76)		1.45 (1.21–1.75)	
high exposure	1.83 (1.53–2.19)		1.82 (1.52–2.18)	
HR per exposure level increase	1.21 (1.15–1.26)		1.20 (1.14–1.25)	
CVD		<0.001		<0.001
no exposure	Ref.		Ref.	
low exposure	1.03 (0.97–1.09)		1.03 (0.97–1.08)	
medium exposure	1.06 (1.01–1.12)		1.06 (1.01–1.11)	
high exposure	1.14 (1.08–1.20)		1.14 (1.08–1.19)	
HR per exposure level increase	1.05 (1.03–1.06)		1.05 (1.03–1.06)	

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference.
Main model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.
Model 2d was further adjusted for air pollutants, solid fuel use, and passive smoking.

Table S17. Associations of total lifetime occupational hazard exposure with transitions from baseline to first cardiopulmonary disease and then to comorbidity

Total lifetime occupational hazard exposure	**Baseline → COPD/CVD**	**COPD/CVD → comorbidity**
	HR (95%CI)	**HR (95%CI)**
no exposure	Ref.	Ref.
low exposure	1.03 (0.97-1.09)	1.10 (0.85-1.37)
medium exposure	1.08 (1.02-1.14)	1.18 (0.94-1.45)
high exposure	1.16 (1.10-1.22)	1.33 (1.06-1.61)
HR per exposure level increase	1.05 (1.04-1.07)	1.10 (1.06-1.14)
P for trend	<0.001	<0.001

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval; Ref., reference.

"Comorbidity" refers to the incidence of both COPD and CVD during follow-up.

Among 85,287 participants, 18,294 transitioned from no disease to one disease, and 848 subsequently progressed to comorbidity.

Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS, genotyping array, and the first 10 principal components of ancestry.

Table S18. Ascertainment of outcome

	Impatient-ICD10 (41270)	*First occurrence-ICD10	Impatient-ICD9 (41271)
COPD	J40-J44	131484-131492	491-492, 496
CVD	I00-I09; I15-I99	131270-131286; 131296-131422	390-398; 406-459

ICD: international classification disease.

* A set of 'first occurrence' data-fields have been generated that map the clinical codes from primary care, hospital inpatient admissions, death records and self-reported medical conditions to 3-character ICD-10 codes and provide, for each participant, the date that code first occurred in any source.

Outcomes were identified using multiple UK Biobank data sources to maximize case ascertainment. Impatient Data: Linked to hospital records from England's Hospital Episode Statistics (HES), Scotland's Scottish Morbidity Records (SMR), and Wales' Patient Episode Database (PEDW), using ICD-10 or ICD-9 codes to capture hospital admissions, both overnight and day cases. First Occurrence Data: A curated dataset consolidating the earliest recorded instance of specific health conditions across multiple sources, including hospital records, death registries, cancer registries, primary care data (~45% of participants), and self-reported conditions. Primary care data were mapped to ICD-10 codes by UK Biobank researchers.

Figure S1. Flowchart of participant enrolment

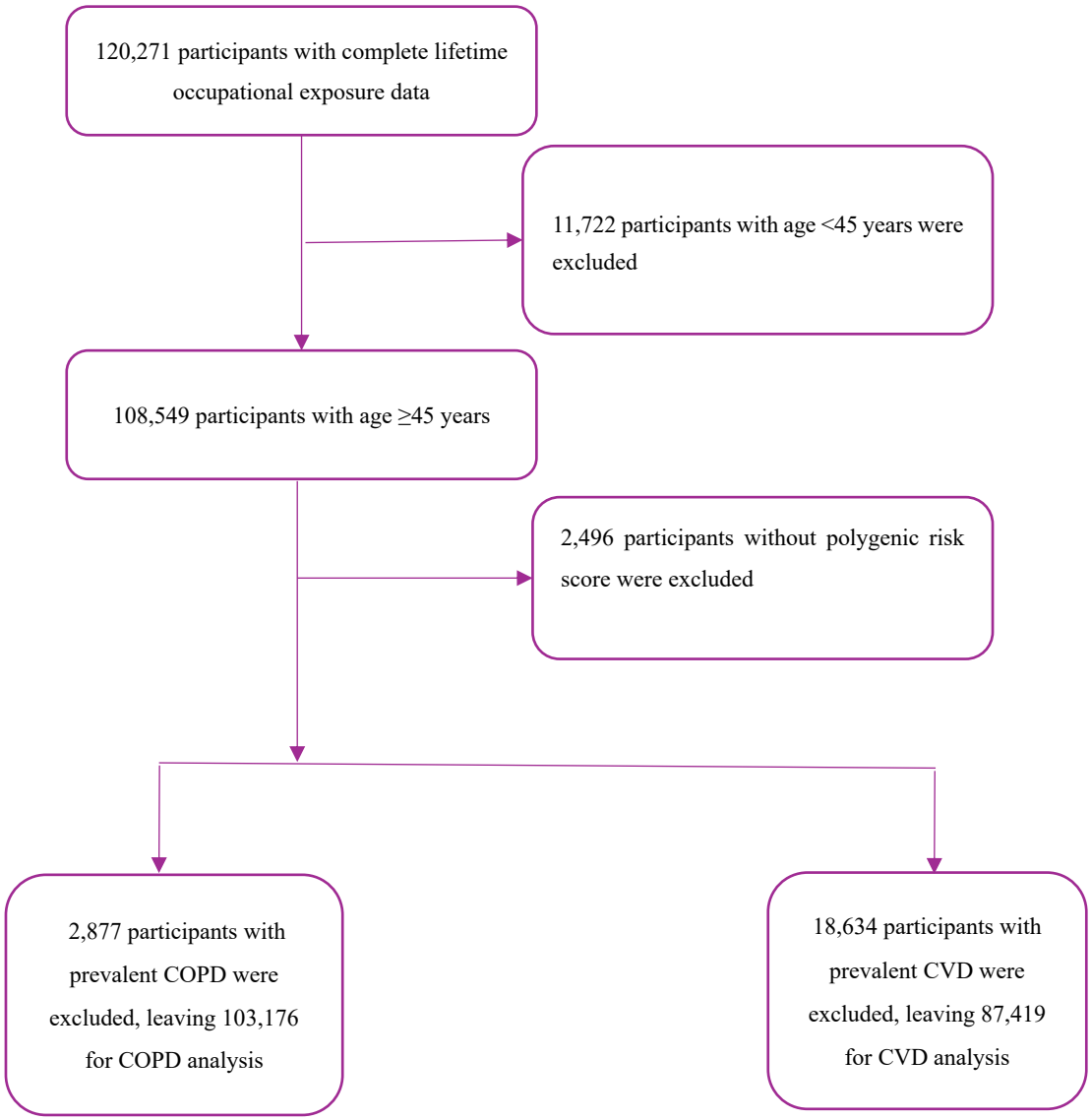
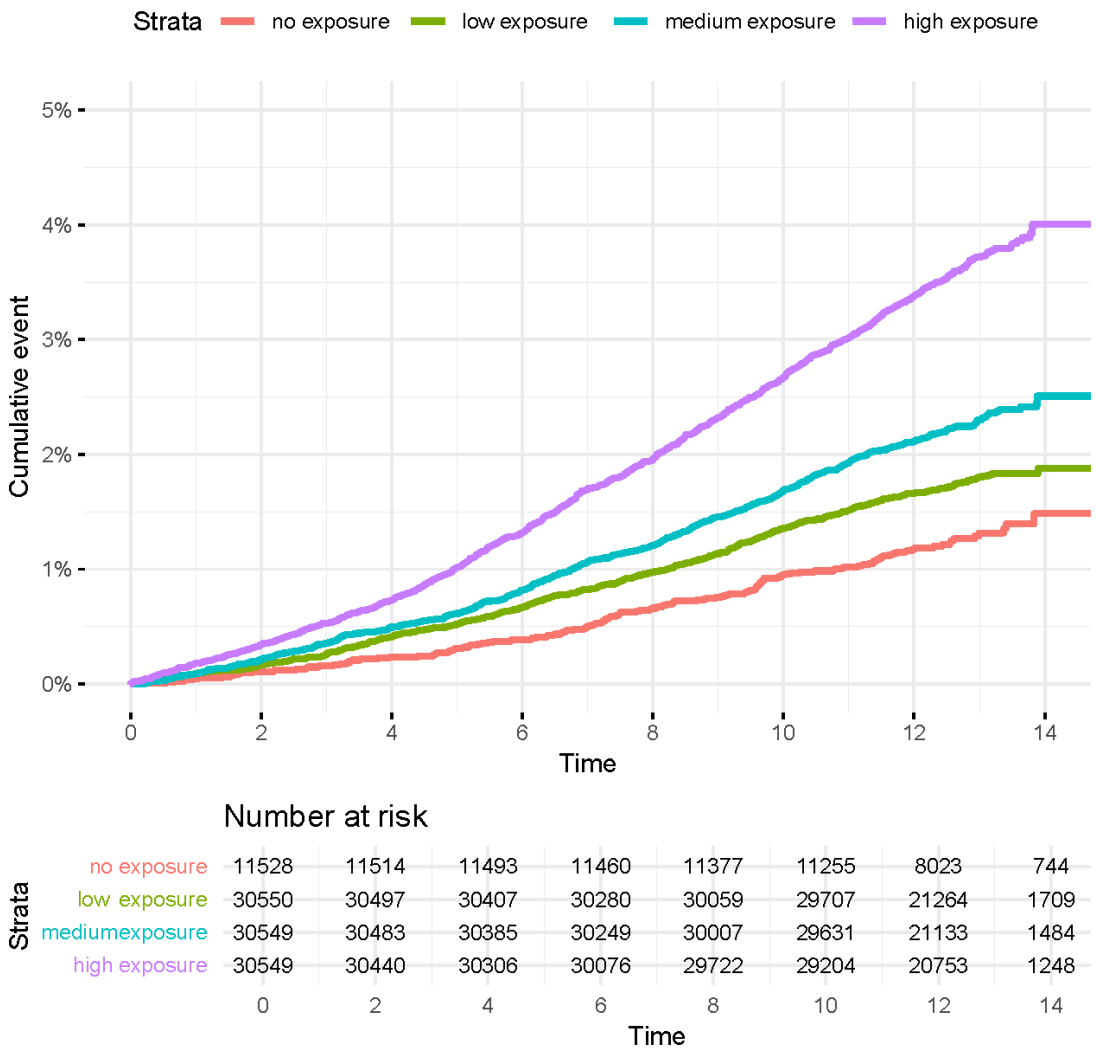


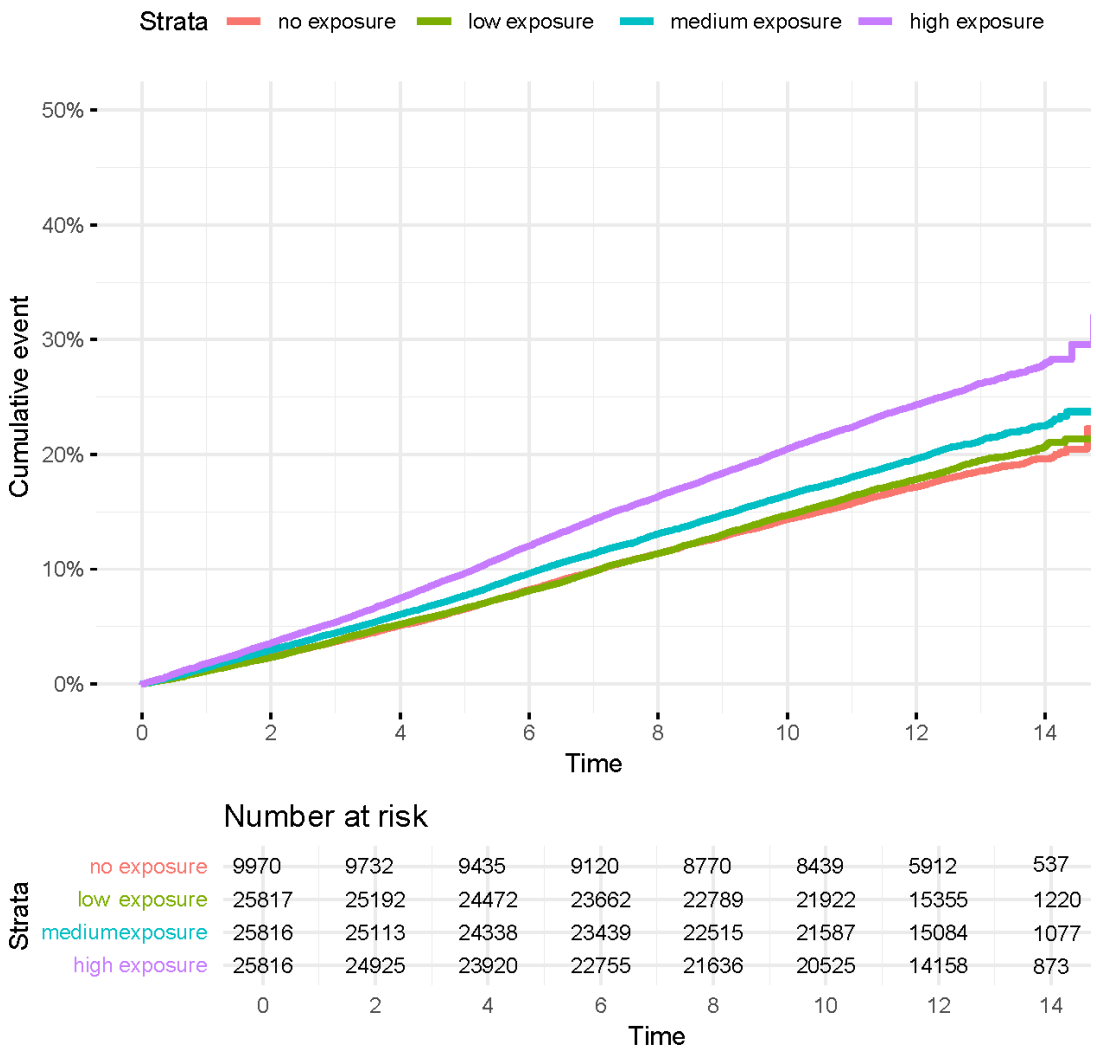
Figure S2. Cumulative risk of COPD stratified by total lifetime occupational hazard exposure



COPD, chronic obstructive pulmonary disease.

Associations of total lifetime occupational hazard exposure with the risk of incident COPD were examined using the Fine–Grey proportional sub-distribution hazards model. Cumulative incidence curves were plotted using the cumulative incidence function as implemented in the 'cmprsk' R package.

Figure S3. Cumulative risk of CVD stratified by total lifetime occupational hazard exposure

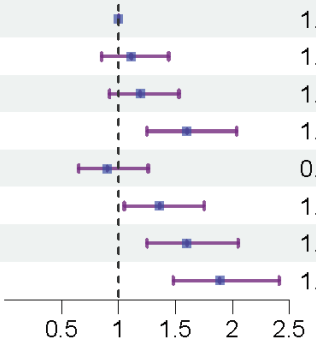


CVD, cardiovascular disease.

Associations of total lifetime occupational hazard exposure with the risk of incident CVD were examined using the Fine–Grey proportional sub-distribution hazards model. Cumulative incidence curves were plotted using the cumulative incidence function as implemented in the 'cmprsk' R package.

Figure S4. Joint associations of total lifetime occupational hazard exposure and COPD PRS on incident COPD

Total lifetime occupational hazard exposure	COPD PRS	Cases	Total number	Person year	HR(95% CI)	p for interaction
no exposure	low	75	5795	72415	1.00 (1.00 to 1.00)	<0.001
low expousre	low	237	15253	189880	1.11 (0.85 to 1.44)	
medium exposure	low	289	15316	190395	1.19 (0.92 to 1.53)	
high exposure	low	500	15224	187669	1.60 (1.25 to 2.04)	
no exposure	high	68	5733	71537	0.90 (0.65 to 1.26)	
low expousre	high	288	15297	190221	1.36 (1.05 to 1.75)	
medium exposure	high	389	15233	188833	1.60 (1.25 to 2.05)	
high exposure	high	580	15325	188511	1.89 (1.48 to 2.41)	

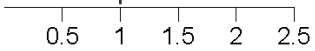


COPD, chronic obstructive pulmonary disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval.

Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, genotyping array, and the first 10 principal components of ancestry.

Figure S5. Joint associations of total lifetime occupational hazard exposure and CVD PRS on incident CVD

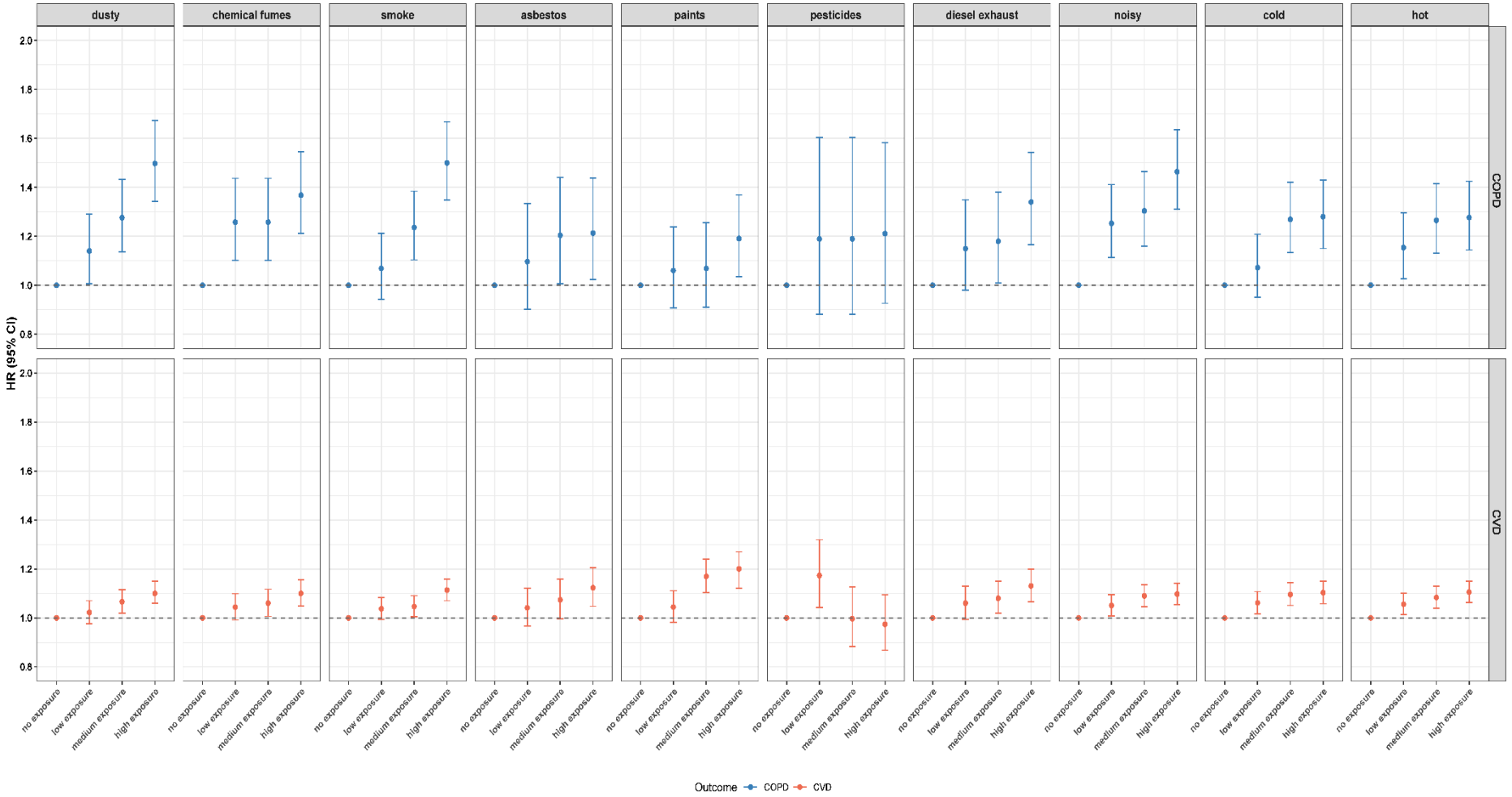
Total lifetime occupational hazard exposure	CVD PRS	Cases	Total number	Person year	HR(95% CI)	p for interaction
no exposure	low	868	5136	59406	1.00 (1.00 to 1.00)	<0.001
low expousre	low	2229	12963	149339	1.06 (0.98 to 1.15)	
medium exposure	low	2436	12880	146850	1.10 (1.02 to 1.19)	
high exposure	low	2947	12731	141496	1.18 (1.09 to 1.27)	
no exposure	high	965	4834	54866	1.23 (1.12 to 1.35)	
low expousre	high	2479	12854	146673	1.24 (1.14 to 1.36)	
medium exposure	high	2708	12936	146428	1.27 (1.17 to 1.37)	
high exposure	high	3403	13085	143215	1.36 (1.26 to 1.47)	



CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval.

Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, genotyping array, and the first 10 principal components of ancestry.

Figure S6. Associations between lifetime exposure to specific occupational hazard and the risk of COPD and CVD (using 0.5-year adjustment for job duration)



COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PRS, polygenic risk score; HR, hazard ratios; CI, confidence interval.
Model was adjusted for age at recruitment, sex, assessment center (strata), body mass index, ethnicity, Townsend Deprivation Index, household income, education, smoking status, alcohol drinking frequency, physical activity, healthy diet score, and sleep duration, PRS (COPD PRS, CVD PRS, respectively), genotyping array, and the first 10 principal components of ancestry.

Additional file 2

Content

Table S1. Mean concentration of biomarkers

Biomarkers#	Field id	Unit	COPD (N=79753)		CVD (n=67583)	
			Mean	SD	Mean	SD
Inflammatory-related biomarkers						
WBC count	30000	10^9 cells/L	6.668	2.004	6.655	2.016
Lymphocyte count	30120	10^9 cells/L	1.923	1.211	1.924	1.231
Lymphocyte percentage	30180	%	29.116	7.327	29.195	7.294
Neutrophil count	30130	10^9 cells/L	4.069	1.301	4.058	1.299
Neutrophil percentage	30190	%	60.580	8.337	60.534	8.305
Monocyte count	30140	10^9 cells/L	0.468	0.195	0.465	0.195
Monocyte percentage	30200	%	7.165	2.660	7.136	2.658
Basophill count	30160	10^9 cells/L	0.032	0.047	0.032	0.045
Basophill percentage	30220	%	0.560	0.558	0.559	0.548
C reactive protein	30710	mg/L	2.259	3.949	2.236	3.892
Platelet count	30080	10^9 cells/L	250.344	57.603	251.696	57.269
Erythrocyte-related biomarkers						
Erythrocyte count	30010	10^12 cells/L	4.512	0.399	4.512	0.398
High light scatter reticulocyte count	30300	10^12 cells/L	0.017	0.010	0.017	0.010
Reticulocyte count	30250	10^12 cells/L	0.059	0.041	0.059	0.042
Red blood cell distribution width	30070	%	13.439	0.901	13.429	0.896
Haematocrit percentage	30030	%	41.112	3.424	41.093	3.419
Haemoglobin concentration	30020	g/dL	14.193	1.196	14.188	1.195
Renal function-related biomarkers						
Cystatin C	30720	mg/L	0.892	0.147	0.886	0.135
Urate	30880	μmol/L	306.006	77.866	304.313	76.954
Urea	30670	mmol/L	5.399	1.283	5.361	1.244
Creatinine	30700	umol/L	72.126	15.343	71.725	14.331
Liver function-related biomarkers						
Alanine aminotransferase	30620	U/L	22.876	13.076	22.766	13.081
Alkaline phosphatase	30610	U/L	81.907	23.829	81.662	23.441
Aspartate aminotransferase	30650	U/L	25.878	8.886	25.732	8.823
Gamma-glutamyl transferase	30730	U/L	34.400	33.927	33.971	33.260

Total bilirubin	30840	μmol/L	9.283	4.482	9.218	4.412
Total protein	30860	g/L	72.198	3.970	72.237	3.949
Albumin	30600	g/L	45.311	2.539	45.345	2.528
Circulating metabolic biomarkers						
Glycated haemoglobin (HbA1c)	mmol/mol	30750	35.515	5.514	35.374	5.375
Glucose	mmol/L	30740	5.075	1.051	5.064	1.022
Cholesterol	mmol/L	30690	5.773	1.101	5.837	1.073
Triglycerides	mmol/L	30870	1.680	0.935	1.678	0.933
LDL direct	mmol/L	30780	3.612	0.841	3.659	0.825
HDL cholesterol	mmol/L	30760	1.484	0.374	1.494	0.372
Apolipoprotein A	g/L	30630	1.563	0.269	1.569	0.268
Apolipoprotein B	g/L	30640	1.040	0.232	1.050	0.230

COPD: chronic obstructive pulmonary disease; CVD: cardiovascular disease; SD: standard deviation; WBC: white blood cell; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Table S2. Associations of total lifetime occupational hazard exposure with biomarkers after removing the baseline COPD cases (n=79753)

Biomarkers[#]	Beta[§]	95% CI	P_{FDR}-value
Inflammatory-related biomarkers			
WBC count	0.014	(0.006, 0.021)	<0.001*
Lymphocyte count	0.007	(0.000, 0.014)	0.200
Lymphocyte percentage	-0.001	(-0.009, 0.006)	0.817
Neutrophil count	0.012	(0.005, 0.019)	0.009*
Neutrophil percentage	0.001	(-0.007, 0.008)	0.903
Monocyte count	0.008	(0.001, 0.015)	0.148
Monocyte percentage	-0.001	(-0.009, 0.006)	0.805
Basophill count	0.007	(-0.001, 0.014)	0.207
Basophill percentage	0.004	(-0.004, 0.011)	0.495
C reactive protein	0.011	(0.004, 0.018)	0.019*
Platelet count	0.010	(0.002, 0.017)	0.048*
Erythrocyte-related biomarkers			
Erythrocyte count	-0.004	(-0.010, 0.002)	0.382
High light scatter reticulocyte count	0.007	(0.000, 0.014)	0.198
Reticulocyte count	0.002	(-0.005, 0.009)	0.720
Red blood cell distribution width	-0.002	(-0.009, 0.006)	0.782
Haematocrit percentage	-0.005	(-0.011, 0.001)	0.207
Haemoglobin concentration	-0.003	(-0.009, 0.003)	0.494
Renal function-related biomarkers			
Cystatin C	0.000	(-0.007, 0.007)	0.994
Urate	0.001	(-0.004, 0.007)	0.782
Urea	0.003	(-0.004, 0.010)	0.630
Creatinine	-0.003	(-0.009, 0.003)	0.527
Liver function-related biomarkers			
Alanine aminotransferase	0.004	(-0.003, 0.011)	0.470
Alkaline phosphatase	0.012	(0.005, 0.019)	0.009*
Aspartate aminotransferase	0.003	(-0.004, 0.011)	0.537
Gamma-glutamyl transferase	0.005	(-0.002, 0.012)	0.382
Total bilirubin	-0.011	(-0.018, -0.004)	0.019*
Total protein	0.003	(-0.004, 0.010)	0.599
Albumin	-0.005	(-0.012, 0.003)	0.398
Circulating metabolic biomarkers			
Glycated haemoglobin	0.009	(0.002, 0.016)	0.069
Glucose	0.002	(-0.005, 0.009)	0.720
Cholesterol	-0.007	(-0.014, 0.001)	0.202
Triglycerides	-0.006	(-0.013, 0.001)	0.224
LDL	-0.007	(-0.014, 0.000)	0.198
HDL	0.000	(-0.006, 0.006)	0.961

Apolipoprotein A	0.001	(-0.005, 0.007)	0.858
Apolipoprotein B	-0.005	(-0.012, 0.002)	0.382

COPD: chronic obstructive pulmonary disease; CI: confidential interval; FDR, false discovery rate; WBC: white blood cell; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Biomarkers concentrations were standardized by z-score.

* indicates FDR-adjusted p values < 0.05.

§Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.

Table S3. Associations of biomarkers with COPD risk (n=79753)

Biomarkers[#]	Beta[§]	95% CI	HR (95%CI)[§]	P_{FDR}-value
Inflammatory-related biomarkers				
WBC count	0.054	(0.042, 0.066)	1.056 (1.043, 1.068)	<0.001*
Lymphocyte count	0.023	(0.006, 0.041)	1.024 (1.006, 1.042)	0.016*
Lymphocyte percentage	-0.151	(-0.199, -0.104)	0.860 (0.820, 0.902)	<0.001*
Neutrophil count	0.199	(0.161, 0.237)	1.220 (1.175, 1.268)	<0.001*
Neutrophil percentage	0.083	(0.036, 0.129)	1.086 (1.037, 1.138)	<0.001*
Monocyte count	0.086	(0.063, 0.108)	1.089 (1.065, 1.115)	<0.001*
Monocyte percentage	0.025	(-0.016, 0.066)	1.025 (0.984, 1.068)	0.285
Basophill count	0.066	(0.041, 0.091)	1.068 (1.042, 1.095)	<0.001*
Basophill percentage	0.049	(0.017, 0.082)	1.050 (1.017, 1.085)	0.006*
C reactive protein	0.117	(0.091, 0.143)	1.124 (1.095, 1.154)	<0.001*
Platelet count	0.084	(0.039, 0.130)	1.088 (1.039, 1.139)	<0.001*
Erythrocyte-related biomarkers				
Erythrocyte count	-0.027	(-0.078, 0.025)	0.974 (0.925, 1.026)	0.375
High light scatter reticulocyte count	0.063	(0.037, 0.090)	1.065 (1.037, 1.094)	<0.001*
Reticulocyte count	0.028	(0.002, 0.053)	1.028 (1.002, 1.054)	0.049*
Red blood cell distribution width	0.125	(0.088, 0.162)	1.133 (1.092, 1.176)	<0.001*
Haematocrit percentage	0.006	(-0.049, 0.060)	1.006 (0.952, 1.062)	0.839
Haemoglobin concentration	-0.027	(-0.083, 0.030)	0.974 (0.920, 1.030)	0.410
Renal function-related biomarkers				
Cystatin C	0.116	(0.091, 0.141)	1.123 (1.095, 1.152)	<0.001*
Urate	0.090	(0.035, 0.144)	1.094 (1.036, 1.155)	0.002*
Urea	0.056	(0.015, 0.097)	1.058 (1.015, 1.102)	0.013*
Creatinine	0.036	(-0.009, 0.082)	1.037 (0.991, 1.086)	0.156
Liver function-related biomarkers				
Alanine aminotransferase	-0.039	(-0.090, 0.012)	0.962 (0.914, 1.012)	0.175
Alkaline phosphatase	0.069	(0.032, 0.107)	1.072 (1.032, 1.113)	<0.001*
Aspartate aminotransferase	-0.014	(-0.060, 0.032)	0.986 (0.941, 1.033)	0.585
Gamma-glutamyl transferase	0.052	(0.017, 0.087)	1.053 (1.017, 1.091)	0.007*
Total bilirubin	-0.097	(-0.152, -0.042)	0.907 (0.859, 0.959)	<0.001*
Total protein	-0.037	(-0.082, 0.009)	0.964 (0.921, 1.009)	0.156
Albumin	-0.115	(-0.162, -0.069)	0.891 (0.851, 0.933)	<0.001*
Circulating metabolic biomarkers				
Glycated haemoglobin	0.108	(0.080, 0.137)	1.114 (1.083, 1.147)	<0.001*
Glucose	0.014	(-0.024, 0.052)	1.014 (0.977, 1.053)	0.520
Cholesterol	-0.147	(-0.193, -0.101)	0.863 (0.825, 0.904)	<0.001*
Triglycerides	-0.047	(-0.094, 0.000)	0.954 (0.910, 1.000)	0.074
LDL	-0.157	(-0.202, -0.111)	0.855 (0.817, 0.895)	<0.001*
HDL	-0.018	(-0.075, 0.039)	0.982 (0.928, 1.040)	0.585

Apolipoprotein A	-0.008	(-0.062, 0.046)	0.992 (0.940, 1.047)	0.792
Apolipoprotein B	-0.129	(-0.175, -0.084)	0.879 (0.839, 0.920)	<0.001*

COPD: chronic obstructive pulmonary disease; HR: hazard ratio; CI: confidential interval; FDR, false discovery rate; WBC: white blood cell; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Biomarkers concentrations were standardized by z-score.

* indicates FDR-adjusted p values < 0.05.

§Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.

Table S4. Mediation analysis of biomarkers as mediators in the relationship between total lifetime occupational hazard exposure and COPD risk (n=79753)

Biomarkers[#]	Total effect			Nature direct effect			Nature indirect effect (Mediation effect)			Mediated Proportion (%)[a]		P for mediated effect
	HR	95% CI	P_{FDR}-value	HR	95% CI	P_{FDR}-value	HR	95% CI	P_{FDR}-value	Proportion (%)	95% CI	
Inflammatory-related biomarkers												
WBC count	1.196	(1.134,1.249)	<0.001*	1.194	(1.132,1.247)	<0.001*	1.001	(1.001,1.002)	<0.001*	0.75	(0.33,1.52)	<0.001*
Lymphocyte count	1.196	(1.132,1.249)	<0.001*	1.195	(1.132,1.249)	<0.001*	1.000	(1.000,1.001)	0.699	0.17	(-0.24,0.42)	0.330
Lymphocyte percentage	1.195	(1.141,1.255)	<0.001*	1.195	(1.140,1.254)	<0.001*	1.000	(0.999,1.002)	0.983	0.15	(-0.75,1.09)	0.800
Neutrophil count	1.196	(1.139,1.250)	<0.001*	1.192	(1.135,1.245)	<0.001*	1.003	(1.001,1.005)	0.010*	1.70	(0.63,2.98)	0.010*
Neutrophil percentage	1.196	(1.137,1.256)	<0.001*	1.196	(1.138,1.256)	<0.001*	1.000	(0.999,1.001)	0.990	0.04	(-0.54,0.52)	0.990
Monocyte count	1.197	(1.143,1.262)	<0.001*	1.196	(1.142,1.261)	<0.001*	1.001	(1.000,1.002)	0.148	0.43	(0.04,0.93)	0.020*
Monocyte percentage	1.195	(1.135,1.252)	<0.001*	1.195	(1.136,1.252)	<0.001*	1.000	(1.000,1.000)	0.983	-0.04	(-0.31,0.18)	0.780
Basophill count	1.196	(1.134,1.271)	<0.001*	1.195	(1.133,1.271)	<0.001*	1.000	(1.000,1.001)	0.555	0.26	(-0.03,0.63)	0.120
Basophill percentage	1.195	(1.142,1.257)	<0.001*	1.195	(1.142,1.257)	<0.001*	1.000	(1.000,1.001)	0.694	0.13	(-0.10,0.53)	0.280
C reactive protein	1.192	(1.142,1.248)	<0.001*	1.191	(1.142,1.247)	<0.001*	1.001	(1.000,1.002)	<0.001*	0.74	(0.20,1.29)	<0.001*
Platelet count	1.191	(1.137,1.245)	<0.001*	1.191	(1.138,1.245)	<0.001*	1.000	(0.999,1.001)	0.885	0.16	(-0.50,0.76)	0.550

Erythrocyte-related biomarkers												
Erythrocyte count	1.196	(1.132,1.260)	<0.001*	1.196	(1.132,1.260)	<0.001*	1.000	(1.000,1.001)	0.721	0.11	(-0.13,0.52)	0.370
High light scatter reticulocyte count	1.195	(1.142,1.248)	<0.001*	1.195	(1.142,1.247)	<0.001*	1.000	(1.000,1.001)	0.572	0.17	(-0.04,0.55)	0.170
Reticulocyte count	1.195	(1.133,1.251)	<0.001*	1.195	(1.134,1.251)	<0.001*	1.000	(0.999,1.000)	0.983	-0.02	(-0.33,0.11)	0.740
Red blood cell distribution width	1.196	(1.131,1.256)	<0.001*	1.196	(1.131,1.256)	<0.001*	1.000	(0.999,1.001)	0.983	-0.11	(-0.56,0.33)	0.640
Haematocrit percentage	1.196	(1.127,1.250)	<0.001*	1.196	(1.127,1.249)	<0.001*	1.000	(0.999,1.001)	0.987	0.01	(-0.37,0.37)	0.910
Haemoglobin concentration	1.195	(1.132,1.260)	<0.001*	1.195	(1.132,1.260)	<0.001*	1.000	(1.000,1.001)	0.983	0.03	(-0.19,0.33)	0.750
Renal function-related biomarkers												
Cystatin C	1.193	(1.133,1.258)	<0.001*	1.193	(1.133,1.259)	<0.001*	1.000	(0.999,1.001)	0.987	0.00	(-0.37,0.53)	0.960
Urate	1.191	(1.135,1.261)	<0.001*	1.191	(1.135,1.260)	<0.001*	1.000	(0.999,1.001)	0.983	0.13	(-0.45,0.72)	0.690
Urea	1.194	(1.133,1.272)	<0.001*	1.194	(1.134,1.272)	<0.001*	1.000	(1.000,1.001)	0.983	0.03	(-0.22,0.32)	0.850
Creatinine	1.195	(1.135,1.249)	<0.001*	1.195	(1.134,1.249)	<0.001*	1.000	(1.000,1.000)	0.983	-0.02	(-0.20,0.21)	0.840
Liver function-related biomarkers												
Alanine aminotransferase	1.189	(1.123,1.255)	<0.001*	1.190	(1.123,1.255)	<0.001*	1.000	(0.998,1.000)	0.694	-0.31	(-1.29,0.29)	0.300

Alkaline phosphatase	1.195	(1.131,1.256)	<0.001*	1.195	(1.132,1.256)	<0.001*	1.001	(1.000,1.001)	0.555	0.32	(-0.14,0.82)	0.150
Aspartate aminotransferase	1.191	(1.124,1.251)	<0.001*	1.192	(1.125,1.252)	<0.001*	1.000	(0.999,1.000)	0.722	-0.18	(-0.66,0.20)	0.390
Gamma-glutamyl transferase	1.195	(1.140,1.243)	<0.001*	1.195	(1.140,1.243)	<0.001*	1.000	(1.000,1.001)	0.885	0.07	(-0.13,0.34)	0.530
Total bilirubin	1.195	(1.140,1.268)	<0.001*	1.195	(1.139,1.266)	<0.001*	1.001	(1.000,1.002)	0.555	0.44	(-0.15,1.27)	0.150
Total protein	1.194	(1.135,1.254)	<0.001*	1.194	(1.135,1.255)	<0.001*	1.000	(0.999,1.000)	0.722	-0.12	(-0.59,0.14)	0.410
Albumin	1.196	(1.136,1.258)	<0.001*	1.195	(1.135,1.258)	<0.001*	1.001	(1.000,1.002)	0.694	0.32	(-0.24,1.02)	0.250
Circulating metabolic biomarkers												
Glycated haemoglobin	1.195	(1.138,1.267)	<0.001*	1.194	(1.137,1.266)	<0.001*	1.001	(1.000,1.002)	0.148	0.57	(0.11,1.22)	0.020*
Glucose	1.194	(1.135,1.256)	<0.001*	1.194	(1.135,1.256)	<0.001*	1.000	(1.000,1.000)	0.983	-0.03	(-0.31,0.13)	0.830
Cholesterol	1.194	(1.133,1.274)	<0.001*	1.193	(1.132,1.274)	<0.001*	1.001	(1.000,1.002)	0.370	0.52	(-0.01,1.15)	0.060
Triglycerides	1.195	(1.141,1.262)	<0.001*	1.194	(1.141,1.262)	<0.001*	1.000	(1.000,1.001)	0.699	0.14	(-0.11,0.55)	0.340
LDL	1.195	(1.140,1.255)	<0.001*	1.194	(1.140,1.253)	<0.001*	1.001	(1.000,1.002)	0.476	0.61	(-0.10,1.36)	0.090
HDL	1.195	(1.140,1.250)	<0.001*	1.195	(1.140,1.250)	<0.001*	1.000	(1.000,1.000)	0.987	0.00	(-0.14,0.14)	0.930
Apolipoprotein A	1.196	(1.133,1.252)	<0.001*	1.196	(1.133,1.252)	<0.001*	1.000	(1.000,1.000)	0.987	0.01	(-0.16,0.21)	0.960

Apolipoprotein B	1.195	(1.133,1.253)	<0.001*	1.194	(1.132,1.253)	<0.001*	1.001	(1.000,1.002)	0.694	0.35	(-0.24,1.08)	0.290

COPD: chronic obstructive pulmonary disease; HR: hazard ratio; CI: confidential interval; FDR, false discovery rate; WBC: white blood cell; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Biomarkers concentrations were standardized by z-score.

* indicates p values < 0.05.

Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.

Table 5. Associations of total life occupational exposure with biomarkers after removing the baseline CVD cases (n=67583)

Biomarkers[#]	Beta[§]	95% CI	P_{FDR}-value
Inflammatory-related biomarkers			
WBC count	0.013	(0.005, 0.021)	0.009*
Lymphocyte count	0.006	(-0.002, 0.014)	0.340
Lymphocyte percentage	-0.003	(-0.011, 0.005)	0.596
Neutrophil count	0.013	(0.005, 0.021)	0.009*
Neutrophil percentage	0.003	(-0.005, 0.011)	0.538
Monocyte count	0.004	(-0.004, 0.012)	0.500
Monocyte percentage	-0.005	(-0.013, 0.003)	0.440
Basophill count	0.006	(-0.002, 0.014)	0.340
Basophill percentage	0.004	(-0.004, 0.012)	0.538
C reactive protein	0.013	(0.005, 0.020)	0.009*
Platelet count	0.007	(-0.001, 0.015)	0.333
Erythrocyte-related biomarkers			
Erythrocyte count	-0.003	(-0.009, 0.004)	0.538
High light scatter reticulocyte count	0.007	(-0.001, 0.014)	0.337
Reticulocyte count	0.000	(-0.008, 0.008)	0.976
Red blood cell distribution width	-0.004	(-0.012, 0.004)	0.506
Haematocrit percentage	-0.003	(-0.010, 0.003)	0.506
Haemoglobin concentration	0.000	(-0.006, 0.006)	0.959
Renal function-related biomarkers			
Cystatin C	0.001	(-0.006, 0.008)	0.848
Urate	0.001	(-0.005, 0.007)	0.826
Urea	0.004	(-0.004, 0.012)	0.500
Creatinine	-0.004	(-0.010, 0.003)	0.461
Liver function-related biomarkers			
Alanine aminotransferase	0.006	(-0.002, 0.013)	0.345
Alkaline phosphatase	0.010	(0.002, 0.017)	0.086
Aspartate aminotransferase	0.003	(-0.005, 0.011)	0.541
Gamma-glutamyl transferase	0.006	(-0.002, 0.013)	0.340
Total bilirubin	-0.012	(-0.020, -0.005)	0.015*
Total protein	0.001	(-0.007, 0.009)	0.891
Albumin	-0.005	(-0.013, 0.003)	0.461
Circulating metabolic biomarkers			
Glycated haemoglobin	0.009	(0.001, 0.017)	0.116
Glucose	0.004	(-0.004, 0.012)	0.506
Cholesterol	-0.006	(-0.014, 0.001)	0.340
Triglycerides	-0.006	(-0.014, 0.001)	0.340
LDL	-0.006	(-0.014, 0.001)	0.340
HDL	-0.001	(-0.008, 0.005)	0.806
Apolipoprotein A	0.000	(-0.007, 0.007)	0.959

Apolipoprotein B	-0.005	(-0.013, 0.003)	0.461

CVD: cardiovascular disease; CI: confidential interval; FDR, false discovery rate; WBC: white blood cell; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Biomarkers concentrations were standardized by z-score.

* indicates FDR-adjusted p values < 0.05.

§Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.

Table S6. Associations of biomarkers with CVD risk (n=67583)

Biomarkers[#]	Beta[§]	95% CI	HR (95%CI)[§]	P_{FDR}-value
Inflammatory-related biomarkers				
WBC count	0.032	(0.021, 0.043)	1.033 (1.021, 1.044)	<0.001*
Lymphocyte count	0.012	(-0.001, 0.024)	1.012 (0.999, 1.024)	0.082
Lymphocyte percentage	-0.043	(-0.060, -0.026)	0.958 (0.942, 0.974)	<0.001*
Neutrophil count	0.046	(0.029, 0.063)	1.047 (1.030, 1.065)	<0.001*
Neutrophil percentage	0.024	(0.007, 0.041)	1.024 (1.007, 1.042)	0.008*
Monocyte count	0.044	(0.031, 0.058)	1.045 (1.031, 1.060)	<0.001*
Monocyte percentage	0.027	(0.012, 0.043)	1.028 (1.012, 1.044)	<0.001*
Basophill count	0.013	(-0.003, 0.030)	1.013 (0.997, 1.031)	0.143
Basophill percentage	0.010	(-0.007, 0.027)	1.010 (0.993, 1.027)	0.299
C reactive protein	0.052	(0.038, 0.066)	1.054 (1.039, 1.068)	<0.001*
Platelet count	0.006	(-0.012, 0.023)	1.006 (0.988, 1.024)	0.555
Erythrocyte-related biomarkers				
Erythrocyte count	-0.039	(-0.058, -0.019)	0.962 (0.943, 0.981)	<0.001*
High light scatter reticulocyte count	0.028	(0.014, 0.042)	1.028 (1.014, 1.043)	<0.001*
Reticulocyte count	0.014	(0.001, 0.027)	1.014 (1.001, 1.027)	0.042*
Red blood cell distribution width	0.062	(0.046, 0.078)	1.064 (1.047, 1.081)	<0.001*
Haematocrit percentage	-0.035	(-0.055, -0.014)	0.966 (0.946, 0.986)	0.002*
Haemoglobin concentration	-0.039	(-0.061, -0.018)	0.961 (0.941, 0.982)	<0.001*
Renal function-related biomarkers				
Cystatin C	0.081	(0.065, 0.096)	1.084 (1.068, 1.101)	<0.001*
Urate	0.050	(0.030, 0.071)	1.052 (1.030, 1.073)	<0.001*
Urea	0.032	(0.015, 0.049)	1.033 (1.016, 1.050)	<0.001*
Creatinine	0.006	(-0.013, 0.026)	1.006 (0.987, 1.026)	0.555
Liver function-related biomarkers				
Alanine aminotransferase	0.002	(-0.016, 0.019)	1.002 (0.985, 1.019)	0.865
Alkaline phosphatase	0.022	(0.005, 0.038)	1.022 (1.005, 1.039)	0.017*
Aspartate aminotransferase	0.021	(0.006, 0.036)	1.021 (1.006, 1.037)	0.010*
Gamma-glutamyl transferase	0.034	(0.019, 0.049)	1.035 (1.019, 1.050)	<0.001*
Total bilirubin	0.001	(-0.017, 0.018)	1.001 (0.983, 1.018)	0.954
Total protein	-0.027	(-0.044, -0.010)	0.973 (0.957, 0.990)	0.004*
Albumin	-0.076	(-0.093, -0.058)	0.927 (0.912, 0.943)	<0.001*
Circulating metabolic biomarkers				
Glycated haemoglobin	0.049	(0.034, 0.063)	1.050 (1.034, 1.065)	<0.001*
Glucose	0.015	(0.000, 0.030)	1.015 (1.000, 1.031)	0.061
Cholesterol	-0.033	(-0.049, -0.016)	0.968 (0.952, 0.984)	<0.001*
Triglycerides	-0.006	(-0.023, 0.011)	0.994 (0.977, 1.011)	0.555
LDL	-0.025	(-0.042, -0.008)	0.975 (0.959, 0.992)	0.005*
HDL	-0.037	(-0.058, -0.017)	0.963 (0.944, 0.983)	<0.001*
Apolipoprotein A	-0.034	(-0.053, -0.014)	0.967 (0.948, 0.986)	0.002*

Apolipoprotein B	-0.011	(-0.028, 0.005)	0.989 (0.972, 1.005)	0.212

CVD: cardiovascular disease; HR: hazard ratio; CI: confidential interval; FDR, false discovery rate; WBC: white blood cell; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Biomarkers concentrations were standardized by z-score.

* indicates FDR-adjusted p values < 0.05.

§Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.

Table 7. Mediation analysis of biomarkers as mediators in the relationship between total lifetime occupational hazard exposure and CVD risk (n=67583)

Biomarkers[#]	Total effect			Nature direct effect			Nature indirect effect (Mediation effect)			Mediated Proportion (%)[a]		P for mediated effect
	HR	95% CI	P_{FDR}-value	HR	95% CI	P_{FDR}-value	HR	95% CI	P_{FDR}-value	Proportion (%)	95% CI	
LOE												
Inflammatory-related biomarkers												
WBC count	1.043	(1.024,1.059)	<0.001*	1.043	(1.024,1.059)	<0.001*	1.000	(1.000,1.001)	<0.001*	1.20	(0.30,2.71)	<0.001*
Lymphocyte count	1.043	(1.024,1.061)	<0.001*	1.043	(1.025,1.061)	<0.001*	1.000	(1.000,1.000)	0.966	0.02	(-0.77,0.51)	0.840
Lymphocyte percentage	1.043	(1.028,1.061)	<0.001*	1.043	(1.027,1.061)	<0.001*	1.000	(1.000,1.001)	0.708	0.40	(-0.72,1.70)	0.440
Neutrophil count	1.043	(1.021,1.058)	<0.001*	1.042	(1.020,1.058)	<0.001*	1.001	(1.000,1.001)	<0.001*	1.74	(0.51,4.08)	<0.001*
Neutrophil percentage	1.043	(1.029,1.061)	<0.001*	1.043	(1.029,1.061)	<0.001*	1.000	(1.000,1.000)	0.708	0.28	(-0.48,1.24)	0.420
Monocyte count	1.043	(1.026,1.063)	<0.001*	1.043	(1.026,1.063)	<0.001*	1.000	(1.000,1.001)	0.708	0.47	(-0.52,1.39)	0.280
Monocyte percentage	1.043	(1.023,1.060)	<0.001*	1.043	(1.023,1.060)	<0.001*	1.000	(1.000,1.000)	0.411	-0.39	(-1.32,0.12)	0.100
Basophill count	1.043	(1.024,1.059)	<0.001*	1.043	(1.024,1.059)	<0.001*	1.000	(1.000,1.000)	0.708	0.18	(-0.30,0.80)	0.370
Basophill percentage	1.043	(1.024,1.063)	<0.001*	1.043	(1.024,1.063)	<0.001*	1.000	(1.000,1.000)	0.938	0.05	(-0.30,0.48)	0.710
C reactive protein	1.043	(1.026,1.065)	<0.001*	1.042	(1.025,1.064)	<0.001*	1.001	(1.000,1.001)	<0.001*	1.78	(0.65,3.36)	<0.001*

Platelet count	1.043	(1.023,1.064)	<0.001*	1.043	(1.022,1.064)	<0.001*	1.000	(1.000,1.000)	0.840	0.14	(-0.36,0.74)	0.590
Erythrocyte-related biomarkers												
Erythrocyte count	1.043	(1.025,1.063)	<0.001*	1.043	(1.025,1.063)	<0.001*	1.000	(1.000,1.000)	0.708	0.20	(-0.32,0.91)	0.400
High light scatter reticulocyte count	1.043	(1.024,1.060)	<0.001*	1.043	(1.024,1.060)	<0.001*	1.000	(1.000,1.000)	0.370	0.39	(-0.05,1.33)	0.070
Reticulocyte count	1.043	(1.026,1.062)	<0.001*	1.043	(1.026,1.062)	<0.001*	1.000	(1.000,1.000)	0.966	0.00	(-0.27,0.36)	0.930
Red blood cell distribution width	1.043	(1.029,1.061)	<0.001*	1.043	(1.029,1.061)	<0.001*	1.000	(0.999,1.000)	0.708	-0.63	(-1.99,0.52)	0.310
Haematocrit percentage	1.043	(1.024,1.060)	<0.001*	1.043	(1.025,1.060)	<0.001*	1.000	(1.000,1.000)	0.708	0.17	(-0.18,0.70)	0.430
Haemoglobin concentration	1.043	(1.026,1.060)	<0.001*	1.043	(1.026,1.060)	<0.001*	1.000	(1.000,1.000)	0.980	0.02	(-0.38,0.37)	0.980
Renal function-related biomarkers												
Cystatin C	1.042	(1.023,1.060)	<0.001*	1.042	(1.023,1.060)	<0.001*	1.000	(0.999,1.001)	0.966	0.27	(-1.81,2.26)	0.790
Urate	1.043	(1.025,1.059)	<0.001*	1.043	(1.025,1.059)	<0.001*	1.000	(1.000,1.001)	0.840	0.18	(-0.84,1.50)	0.590
Urea	1.043	(1.024,1.063)	<0.001*	1.043	(1.024,1.062)	<0.001*	1.000	(1.000,1.001)	0.708	0.37	(-0.54,1.51)	0.320
Creatinine	1.043	(1.029,1.063)	<0.001*	1.043	(1.029,1.063)	<0.001*	1.000	(1.000,1.000)	0.966	0.01	(-0.28,0.31)	0.940
Liver function-related biomarkers												

Alanine aminotransferase	1.043	(1.027,1.063)	<0.001*	1.043	(1.027,1.063)	<0.001*	1.000	(1.000,1.000)	0.966	0.05	(-0.45,0.58)	0.880
Alkaline phosphatase	1.043	(1.024,1.062)	<0.001*	1.042	(1.024,1.061)	<0.001*	1.000	(1.000,1.001)	0.148	0.80	(0.11,1.92)	0.020*
Aspartate aminotransferase	1.043	(1.025,1.061)	<0.001*	1.043	(1.025,1.061)	<0.001*	1.000	(1.000,1.000)	0.708	0.17	(-0.30,0.67)	0.370
Gamma-glutamyl transferase	1.043	(1.025,1.061)	<0.001*	1.043	(1.025,1.061)	<0.001*	1.000	(1.000,1.001)	0.518	0.60	(-0.23,1.63)	0.140
Total bilirubin	1.043	(1.026,1.061)	<0.001*	1.043	(1.026,1.061)	<0.001*	1.000	(1.000,1.000)	0.966	-0.06	(-0.90,0.73)	0.930
Total protein	1.043	(1.025,1.063)	<0.001*	1.043	(1.025,1.063)	<0.001*	1.000	(1.000,1.000)	0.966	-0.06	(-0.70,0.47)	0.760
Albumin	1.043	(1.027,1.058)	<0.001*	1.043	(1.026,1.057)	<0.001*	1.000	(1.000,1.001)	0.586	1.00	(-0.60,2.74)	0.190
Circulating metabolic biomarkers												
Glycated haemoglobin	1.043	(1.026,1.061)	<0.001*	1.042	(1.025,1.061)	<0.001*	1.001	(1.000,1.001)	0.148	1.44	(0.18,2.96)	0.020*
Glucose	1.043	(1.025,1.060)	<0.001*	1.043	(1.025,1.060)	<0.001*	1.000	(1.000,1.000)	0.708	0.16	(-0.20,0.80)	0.370
Cholesterol	1.043	(1.025,1.062)	<0.001*	1.043	(1.025,1.062)	<0.001*	1.000	(1.000,1.001)	0.411	0.51	(-0.01,1.47)	0.090
Triglycerides	1.043	(1.027,1.063)	<0.001*	1.043	(1.027,1.063)	<0.001*	1.000	(1.000,1.000)	0.802	0.11	(-0.40,0.65)	0.520
LDL	1.043	(1.026,1.060)	<0.001*	1.043	(1.026,1.060)	<0.001*	1.000	(1.000,1.001)	0.538	0.43	(-0.12,1.28)	0.160
HDL	1.043	(1.027,1.061)	<0.001*	1.043	(1.027,1.061)	<0.001*	1.000	(1.000,1.000)	0.938	0.12	(-0.48,0.84)	0.690

Apolipoprotein A	1.043	(1.026,1.059)	<0.001*	1.043	(1.026,1.059)	<0.001*	1.000	(1.000,1.000)	0.966	0.02	(-0.51,0.55)	0.870
Apolipoprotein B	1.043	(1.025,1.062)	<0.001*	1.043	(1.026,1.062)	<0.001*	1.000	(1.000,1.000)	0.708	0.14	(-0.19,0.81)	0.440

CVD: cardiovascular disease; HR: hazard ratio; CI: confidential interval; FDR, false discovery rate; WBC: white blood cell; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Biomarkers concentrations were standardized by z-score.

* indicates p values < 0.05.

Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.

Table S8. Mean concentration of metabolites

Field ID	Description	Unit	Subgroup name	COPD (N=23255)		CVD (N=19718)	
				Mean	SD	Mean	SD
23440	Apolipoprotein A1	g/l	Apolipoproteins	1.453	0.238	1.459	0.237
23439	Apolipoprotein B	g/l	Apolipoproteins	0.851	0.195	0.860	0.193
23427	Total Concentration of Lipoprotein Particles	mmol/L	Lipoprotein particle concentrations	0.017	0.002	0.017	0.002
23428	Concentration of VLDL Particles	mmol/L	Lipoprotein particle concentrations	0.000	0.000	0.000	0.000
23429	Concentration of LDL Particles	mmol/L	Lipoprotein particle concentrations	0.001	0.000	0.001	0.000
23430	Concentration of HDL Particles	mmol/L	Lipoprotein particle concentrations	0.015	0.002	0.015	0.002
23431	Average Diameter for VLDL Particles	nm	Lipoprotein particle sizes	38.539	1.235	38.540	1.230
23432	Average Diameter for LDL Particles	nm	Lipoprotein particle sizes	23.921	0.087	23.923	0.087
23433	Average Diameter for HDL Particles	nm	Lipoprotein particle sizes	9.665	0.213	9.668	0.213
23423	Total Lipids in Lipoprotein Particles	mmol/L	Total Lipids	8.845	1.596	8.923	1.573
23424	Total Lipids in VLDL	mmol/L	Total Lipids	2.079	0.832	2.096	0.835
23425	Total Lipids in LDL	mmol/L	Total Lipids	2.489	0.582	2.519	0.574
23426	Total Lipids in HDL	mmol/L	Total Lipids	3.029	0.636	3.045	0.636
23400	Total Cholesterol	mmol/L	Cholesterol	4.645	0.919	4.696	0.899
23403	VLDL Cholesterol	mmol/L	Cholesterol	0.722	0.241	0.731	0.241
23405	LDL Cholesterol	mmol/L	Cholesterol	1.739	0.426	1.761	0.420
23406	HDL Cholesterol	mmol/L	Cholesterol	1.331	0.327	1.340	0.326
23404	Clinical LDL Cholesterol	mmol/L	Cholesterol	2.559	0.713	2.598	0.701
23401	Total Cholesterol Minus HDL-C	mmol/L	Cholesterol	3.314	0.820	3.357	0.807
23402	Remnant Cholesterol (Non-HDL, Non-LDL -Cholesterol)	mmol/L	Cholesterol	1.575	0.408	1.595	0.402
23407	Total Triglycerides	mmol/L	Triglycerides	1.274	0.551	1.278	0.552
23408	Triglycerides in VLDL	mmol/L	Triglycerides	0.890	0.464	0.894	0.465

23409	Triglycerides in LDL	mmol/L	Triglycerides	0.143	0.039	0.143	0.039
23410	Triglycerides in HDL	mmol/L	Triglycerides	0.142	0.045	0.142	0.045
23411	Total Phospholipids in Lipoprotein Particles	mmol/L	Phospholipids	2.927	0.458	2.949	0.450
23412	Phospholipids in VLDL	mmol/L	Phospholipids	0.467	0.175	0.472	0.176
23413	Phospholipids in LDL	mmol/L	Phospholipids	0.607	0.137	0.614	0.135
23414	Phospholipids in HDL	mmol/L	Phospholipids	1.556	0.314	1.563	0.314
23415	Total Esterified Cholesterol	mmol/L	Cholesteryl Esters	3.373	0.662	3.410	0.647
23416	Cholesteryl Esters in VLDL	mmol/L	Cholesteryl Esters	0.432	0.141	0.438	0.140
23417	Cholesteryl Esters in LDL	mmol/L	Cholesteryl Esters	1.271	0.312	1.287	0.309
23418	Cholesteryl Esters in HDL	mmol/L	Cholesteryl Esters	1.035	0.255	1.041	0.255
23419	Total Free Cholesterol	mmol/L	Free Cholesterol	1.272	0.263	1.286	0.258
23420	Free Cholesterol in VLDL	mmol/L	Free Cholesterol	0.290	0.104	0.293	0.105
23421	Free Cholesterol in LDL	mmol/L	Free Cholesterol	0.468	0.118	0.474	0.115
23422	Free Cholesterol in HDL	mmol/L	Free Cholesterol	0.296	0.074	0.298	0.074
23481	Concentration of Chylomicrons and Extremely Large VLDL Particles	mmol/L	Chylomicrons and extremely large VLDL	0.000	0.000	0.000	0.000
23482	Total Lipids in Chylomicrons and Extremely Large VLDL	mmol/L	Chylomicrons and extremely large VLDL	0.219	0.196	0.219	0.197
23483	Phospholipids in Chylomicrons and Extremely Large VLDL	mmol/L	Chylomicrons and extremely large VLDL	0.034	0.031	0.034	0.031
23484	Cholesterol in Chylomicrons and Extremely Large VLDL	mmol/L	Chylomicrons and extremely large VLDL	0.054	0.041	0.054	0.041
23485	Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	mmol/L	Chylomicrons and extremely large VLDL	0.030	0.023	0.030	0.023
23486	Free Cholesterol in Chylomicrons and Extremely Large VLDL	mmol/L	Chylomicrons and extremely large VLDL	0.024	0.019	0.024	0.019
23487	Triglycerides in Chylomicrons and Extremely Large VLDL	mmol/L	Chylomicrons and extremely large VLDL	0.131	0.126	0.131	0.126
23488	Concentration of Very Large VLDL Particles	mmol/L	Very large VLDL	0.000	0.000	0.000	0.000
23489	Total Lipids in Very Large VLDL	mmol/L	Very large VLDL	0.197	0.130	0.199	0.130
23490	Phospholipids in Very Large VLDL	mmol/L	Very large VLDL	0.038	0.025	0.038	0.025
23491	Cholesterol in Very Large VLDL	mmol/L	Very large VLDL	0.053	0.028	0.053	0.028
23492	Cholesteryl Esters in Very Large VLDL	mmol/L	Very large VLDL	0.030	0.014	0.030	0.014

23493	Free Cholesterol in Very Large VLDL	mmol/L	Very large VLDL	0.023	0.014	0.023	0.014
23494	Triglycerides in Very Large VLDL	mmol/L	Very large VLDL	0.107	0.079	0.107	0.079
23495	Concentration of Large VLDL Particles	mmol/L	Large VLDL	0.000	0.000	0.000	0.000
23496	Total Lipids in Large VLDL	mmol/L	Large VLDL	0.323	0.171	0.325	0.172
23497	Phospholipids in Large VLDL	mmol/L	Large VLDL	0.064	0.038	0.065	0.038
23498	Cholesterol in Large VLDL	mmol/L	Large VLDL	0.096	0.046	0.097	0.047
23499	Cholesteryl Esters in Large VLDL	mmol/L	Large VLDL	0.051	0.023	0.051	0.023
23500	Free Cholesterol in Large VLDL	mmol/L	Large VLDL	0.045	0.024	0.046	0.024
23501	Triglycerides in Large VLDL	mmol/L	Large VLDL	0.162	0.089	0.163	0.090
23502	Concentration of Medium VLDL Particles	mmol/L	Medium VLDL	0.000	0.000	0.000	0.000
23503	Total Lipids in Medium VLDL	mmol/L	Medium VLDL	0.570	0.200	0.577	0.200
23504	Phospholipids in Medium VLDL	mmol/L	Medium VLDL	0.129	0.046	0.131	0.046
23505	Cholesterol in Medium VLDL	mmol/L	Medium VLDL	0.174	0.064	0.177	0.064
23506	Cholesteryl Esters in Medium VLDL	mmol/L	Medium VLDL	0.094	0.037	0.096	0.037
23507	Free Cholesterol in Medium VLDL	mmol/L	Medium VLDL	0.080	0.029	0.081	0.029
23508	Triglycerides in Medium VLDL	mmol/L	Medium VLDL	0.268	0.111	0.270	0.112
23509	Concentration of Small VLDL Particles	mmol/L	Small VLDL	0.000	0.000	0.000	0.000
23510	Total Lipids in Small VLDL	mmol/L	Small VLDL	0.408	0.126	0.411	0.127
23511	Phospholipids in Small VLDL	mmol/L	Small VLDL	0.097	0.029	0.098	0.029
23512	Cholesterol in Small VLDL	mmol/L	Small VLDL	0.157	0.050	0.158	0.050
23513	Cholesteryl Esters in Small VLDL	mmol/L	Small VLDL	0.098	0.032	0.099	0.032
23514	Free Cholesterol in Small VLDL	mmol/L	Small VLDL	0.059	0.018	0.060	0.018
23515	Triglycerides in Small VLDL	mmol/L	Small VLDL	0.155	0.058	0.155	0.058
23516	Concentration of Very Small VLDL Particles	mmol/L	Very small VLDL	0.000	0.000	0.000	0.000
23517	Total Lipids in Very Small VLDL	mmol/L	Very small VLDL	0.362	0.086	0.365	0.086
23518	Phospholipids in Very Small VLDL	mmol/L	Very small VLDL	0.106	0.026	0.106	0.026

23519	Cholesterol in Very Small VLDL	mmol/L	Very small VLDL	0.189	0.049	0.191	0.048
23520	Cholesteryl Esters in Very Small VLDL	mmol/L	Very small VLDL	0.130	0.035	0.132	0.034
23521	Free Cholesterol in Very Small VLDL	mmol/L	Very small VLDL	0.059	0.015	0.059	0.014
23522	Triglycerides in Very Small VLDL	mmol/L	Very small VLDL	0.068	0.020	0.068	0.020
23523	Concentration of IDL Particles	mmol/L	IDL	0.000	0.000	0.000	0.000
23524	Total Lipids in IDL	mmol/L	IDL	1.248	0.285	1.263	0.278
23525	Phospholipids in IDL	mmol/L	IDL	0.296	0.065	0.300	0.064
23526	Cholesterol in IDL	mmol/L	IDL	0.853	0.211	0.865	0.206
23527	Cholesteryl Esters in IDL	mmol/L	IDL	0.628	0.157	0.637	0.153
23528	Free Cholesterol in IDL	mmol/L	IDL	0.224	0.055	0.227	0.054
23529	Triglycerides in IDL	mmol/L	IDL	0.099	0.025	0.099	0.025
23530	Concentration of Large LDL Particles	mmol/L	Large LDL	0.001	0.000	0.001	0.000
23531	Total Lipids in Large LDL	mmol/L	Large LDL	1.589	0.368	1.608	0.362
23532	Phospholipids in Large LDL	mmol/L	Large LDL	0.358	0.081	0.363	0.079
23533	Cholesterol in Large LDL	mmol/L	Large LDL	1.135	0.276	1.150	0.272
23534	Cholesteryl Esters in Large LDL	mmol/L	Large LDL	0.836	0.203	0.847	0.200
23535	Free Cholesterol in Large LDL	mmol/L	Large LDL	0.299	0.076	0.303	0.074
23536	Triglycerides in Large LDL	mmol/L	Large LDL	0.096	0.024	0.096	0.024
23537	Concentration of Medium LDL Particles	mmol/L	Medium LDL	0.000	0.000	0.000	0.000
23538	Total Lipids in Medium LDL	mmol/L	Medium LDL	0.616	0.160	0.624	0.159
23539	Phospholipids in Medium LDL	mmol/L	Medium LDL	0.160	0.041	0.162	0.040
23540	Cholesterol in Medium LDL	mmol/L	Medium LDL	0.424	0.114	0.429	0.113
23541	Cholesteryl Esters in Medium LDL	mmol/L	Medium LDL	0.304	0.085	0.308	0.084
23542	Free Cholesterol in Medium LDL	mmol/L	Medium LDL	0.119	0.031	0.121	0.031
23543	Triglycerides in Medium LDL	mmol/L	Medium LDL	0.032	0.010	0.032	0.010
23544	Concentration of Small LDL Particles	mmol/L	Small LDL	0.000	0.000	0.000	0.000

23545	Total Lipids in Small LDL	mmol/L	Small LDL	0.284	0.063	0.287	0.062
23546	Phospholipids in Small LDL	mmol/L	Small LDL	0.088	0.018	0.089	0.018
23547	Cholesterol in Small LDL	mmol/L	Small LDL	0.181	0.043	0.183	0.042
23548	Cholesteryl Esters in Small LDL	mmol/L	Small LDL	0.131	0.032	0.132	0.032
23549	Free Cholesterol in Small LDL	mmol/L	Small LDL	0.050	0.012	0.050	0.012
23550	Triglycerides in Small LDL	mmol/L	Small LDL	0.015	0.005	0.015	0.005
23551	Concentration of Very Large HDL Particles	mmol/L	Very large HDL	0.000	0.000	0.000	0.000
23552	Total Lipids in Very Large HDL	mmol/L	Very large HDL	0.177	0.083	0.179	0.084
23553	Phospholipids in Very Large HDL	mmol/L	Very large HDL	0.084	0.047	0.085	0.047
23554	Cholesterol in Very Large HDL	mmol/L	Very large HDL	0.086	0.036	0.087	0.036
23555	Cholesteryl Esters in Very Large HDL	mmol/L	Very large HDL	0.062	0.029	0.062	0.029
23556	Free Cholesterol in Very Large HDL	mmol/L	Very large HDL	0.024	0.008	0.025	0.008
23557	Triglycerides in Very Large HDL	mmol/L	Very large HDL	0.007	0.003	0.007	0.003
23558	Concentration of Large HDL Particles	mmol/L	Large HDL	0.001	0.001	0.001	0.001
23559	Total Lipids in Large HDL	mmol/L	Large HDL	0.678	0.333	0.684	0.334
23560	Phospholipids in Large HDL	mmol/L	Large HDL	0.335	0.156	0.338	0.156
23561	Cholesterol in Large HDL	mmol/L	Large HDL	0.312	0.173	0.316	0.173
23562	Cholesteryl Esters in Large HDL	mmol/L	Large HDL	0.241	0.134	0.244	0.135
23563	Free Cholesterol in Large HDL	mmol/L	Large HDL	0.071	0.039	0.072	0.039
23564	Triglycerides in Large HDL	mmol/L	Large HDL	0.031	0.012	0.031	0.012
23565	Concentration of Medium HDL Particles	mmol/L	Medium HDL	0.004	0.001	0.004	0.001
23566	Total Lipids in Medium HDL	mmol/L	Medium HDL	1.032	0.216	1.036	0.216
23567	Phospholipids in Medium HDL	mmol/L	Medium HDL	0.484	0.096	0.486	0.096
23568	Cholesterol in Medium HDL	mmol/L	Medium HDL	0.495	0.119	0.497	0.119
23569	Cholesteryl Esters in Medium HDL	mmol/L	Medium HDL	0.407	0.096	0.409	0.096
23570	Free Cholesterol in Medium HDL	mmol/L	Medium HDL	0.087	0.024	0.088	0.024

23571	Triglycerides in Medium HDL	mmol/L	Medium HDL	0.053	0.018	0.053	0.018
23572	Concentration of Small HDL Particles	mmol/L	Small HDL	0.010	0.001	0.010	0.001
23573	Total Lipids in Small HDL	mmol/L	Small HDL	1.142	0.150	1.145	0.150
23574	Phospholipids in Small HDL	mmol/L	Small HDL	0.653	0.088	0.654	0.088
23575	Cholesterol in Small HDL	mmol/L	Small HDL	0.438	0.058	0.440	0.058
23576	Cholesteryl Esters in Small HDL	mmol/L	Small HDL	0.325	0.045	0.326	0.045
23577	Free Cholesterol in Small HDL	mmol/L	Small HDL	0.114	0.015	0.114	0.015
23578	Triglycerides in Small HDL	mmol/L	Small HDL	0.051	0.017	0.051	0.017
23442	Total Fatty Acids	mmol/L	Fatty acids	11.838	2.287	11.905	2.277
23446	Polyunsaturated Fatty Acids	mmol/L	Fatty acids	5.027	0.785	5.058	0.778
23447	Monounsaturated Fatty Acids	mmol/L	Fatty acids	2.776	0.773	2.788	0.772
23448	Saturated Fatty Acids	mmol/L	Fatty acids	4.034	0.902	4.058	0.899
23449	Linoleic Acid	mmol/L	Fatty acids	3.454	0.673	3.488	0.663
23450	Docosahexaenoic Acid	mmol/L	Fatty acids	0.245	0.084	0.246	0.084
23444	Omega-3 Fatty Acids	mmol/L	Fatty acids	0.546	0.220	0.547	0.220
23445	Omega-6 Fatty Acids	mmol/L	Fatty acids	4.480	0.666	4.511	0.658
23443	Degree of Unsaturation	degree	Fatty acids	1.369	0.079	1.370	0.078
23436	Total Cholines	mmol/L	Other lipids	2.557	0.400	2.574	0.394
23437	Phosphatidylcholines	mmol/L	Other lipids	2.103	0.363	2.118	0.359
23438	Sphingomyelins	mmol/L	Other lipids	0.448	0.071	0.451	0.070
23434	Phosphoglycerides	mmol/L	Other lipids	2.273	0.385	2.288	0.381
23460	Alanine	mmol/L	Amino acids	0.301	0.076	0.300	0.076
23461	Glutamine	mmol/L	Amino acids	0.532	0.077	0.531	0.077
23462	Glycine	mmol/L	Amino acids	0.166	0.064	0.167	0.065
23463	Histidine	mmol/L	Amino acids	0.065	0.010	0.065	0.010

23464	Total Concentration of Branched-Chain Amino Acids (Leucine + Isoleucine + Valine)	mmol/L	Amino acids	0.355	0.083	0.354	0.083
23465	Isoleucine	mmol/L	Amino acids	0.050	0.017	0.050	0.017
23466	Leucine	mmol/L	Amino acids	0.101	0.028	0.101	0.028
23467	Valine	mmol/L	Amino acids	0.204	0.041	0.204	0.041
23468	Phenylalanine	mmol/L	Amino acids	0.045	0.010	0.045	0.010
23469	Tyrosine	mmol/L	Amino acids	0.062	0.014	0.062	0.014
23470	Glucose	mmol/L	Glycolysis related metabolites	3.521	0.991	3.513	0.984
23471	Lactate	mmol/L	Glycolysis related metabolites	3.803	1.077	3.813	1.081
23472	Pyruvate	mmol/L	Glycolysis related metabolites	0.079	0.030	0.079	0.030
23473	Citrate	mmol/L	Glycolysis related metabolites	0.063	0.012	0.063	0.012
23477	Acetone	mmol/L	Ketone bodies	0.014	0.005	0.014	0.005
23476	Acetoacetate	mmol/L	Ketone bodies	0.012	0.011	0.012	0.011
23475	Acetate	mmol/L	Ketone bodies	0.015	0.011	0.015	0.011
23474	3-Hydroxybutyrate	mmol/L	Ketone bodies	0.056	0.056	0.056	0.057
23480	Glycoprotein Acetyls	mmol/L	Inflammation	0.779	0.112	0.779	0.112
23478	Creatinine	mmol/L	Fluid balance	0.066	0.013	0.065	0.012
23479	Albumin	g/l	Fluid balance	39.162	3.211	39.194	3.211

COPD: chronic obstructive pulmonary disease; CVD: cardiovascular disease; SD: standard deviation; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Table S9. Associations of total lifetime occupational hazard exposure with metabolites after removing baseline COPD cases (n=23255)

Metabolites#	beta [a]	95% CI	P_{FDR}-value
Apolipoprotein A1	0.003	(-0.009 ,0.015)	0.683
Apolipoprotein B	-0.013	(-0.026 ,0.001)	0.243
Total Concentration of Lipoprotein Particles	0.003	(-0.010 ,0.015)	0.718
Concentration of VLDL Particles	-0.012	(-0.026 ,0.001)	0.243
Concentration of LDL Particles	-0.014	(-0.028 ,-0.001)	0.243
Concentration of HDL Particles	0.005	(-0.007 ,0.017)	0.516
Average Diameter for VLDL Particles	-0.009	(-0.021 ,0.003)	0.277
Average Diameter for LDL Particles	-0.003	(-0.016 ,0.010)	0.730
Average Diameter for HDL Particles	-0.008	(-0.020 ,0.003)	0.267
Total Lipids in Lipoprotein Particles	-0.011	(-0.025 ,0.002)	0.243
Total Lipids in VLDL	-0.012	(-0.025 ,0.001)	0.243
Total Lipids in LDL	-0.011	(-0.025 ,0.002)	0.247
Total Lipids in HDL	0.001	(-0.010 ,0.013)	0.864
Total Cholesterol	-0.010	(-0.023 ,0.003)	0.266
VLDL Cholesterol	-0.012	(-0.025 ,0.001)	0.243
LDL Cholesterol	-0.011	(-0.025 ,0.002)	0.243
HDL Cholesterol	0.000	(-0.011 ,0.012)	0.953
Clinical LDL Cholesterol	-0.009	(-0.023 ,0.004)	0.288
Total Cholesterol Minus HDL-C	-0.011	(-0.025 ,0.002)	0.243
Remnant Cholesterol (Non-HDL, Non-LDL -Cholesterol)	-0.011	(-0.024 ,0.002)	0.247
Total Triglycerides	-0.010	(-0.023 ,0.003)	0.254
Triglycerides in VLDL	-0.011	(-0.023 ,0.002)	0.243
Triglycerides in LDL	-0.007	(-0.020 ,0.006)	0.378

Triglycerides in HDL	-0.006	(-0.019 ,0.008)	0.489
Total Phospholipids in Lipoprotein Particles	-0.007	(-0.020 ,0.006)	0.372
Phospholipids in VLDL	-0.012	(-0.025 ,0.001)	0.243
Phospholipids in LDL	-0.009	(-0.023 ,0.004)	0.288
Phospholipids in HDL	0.003	(-0.009 ,0.014)	0.699
Total Esterified Cholesterol	-0.009	(-0.022 ,0.004)	0.282
Cholesteryl Esters in VLDL	-0.012	(-0.025 ,0.002)	0.243
Cholesteryl Esters in LDL	-0.012	(-0.025 ,0.002)	0.243
Cholesteryl Esters in HDL	0.001	(-0.010 ,0.013)	0.842
Total Free Cholesterol	-0.012	(-0.025 ,0.001)	0.243
Free Cholesterol in VLDL	-0.012	(-0.025 ,0.001)	0.243
Free Cholesterol in LDL	-0.010	(-0.023 ,0.004)	0.274
Free Cholesterol in HDL	-0.003	(-0.014 ,0.008)	0.670
Concentration of Chylomicrons and Extremely Large VLDL Particles	-0.007	(-0.020 ,0.005)	0.372
Total Lipids in Chylomicrons and Extremely Large VLDL	-0.007	(-0.019 ,0.006)	0.378
Phospholipids in Chylomicrons and Extremely Large VLDL	-0.006	(-0.019 ,0.006)	0.411
Cholesterol in Chylomicrons and Extremely Large VLDL	-0.007	(-0.019 ,0.006)	0.399
Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	-0.007	(-0.019 ,0.006)	0.399
Free Cholesterol in Chylomicrons and Extremely Large VLDL	-0.007	(-0.019 ,0.006)	0.399
Triglycerides in Chylomicrons and Extremely Large VLDL	-0.007	(-0.020 ,0.006)	0.372
Concentration of Very Large VLDL Particles	-0.009	(-0.022 ,0.003)	0.267
Total Lipids in Very Large VLDL	-0.010	(-0.022 ,0.003)	0.267
Phospholipids in Very Large VLDL	-0.010	(-0.022 ,0.003)	0.263
Cholesterol in Very Large VLDL	-0.011	(-0.023 ,0.002)	0.247

Cholesteryl Esters in Very Large VLDL	-0.010	(-0.023 ,0.002)	0.251
Free Cholesterol in Very Large VLDL	-0.010	(-0.023 ,0.002)	0.247
Triglycerides in Very Large VLDL	-0.009	(-0.021 ,0.004)	0.279
Concentration of Large VLDL Particles	-0.011	(-0.024 ,0.001)	0.243
Total Lipids in Large VLDL	-0.012	(-0.024 ,0.001)	0.243
Phospholipids in Large VLDL	-0.010	(-0.023 ,0.002)	0.247
Cholesterol in Large VLDL	-0.012	(-0.024 ,0.001)	0.243
Cholesteryl Esters in Large VLDL	-0.012	(-0.025 ,0.001)	0.243
Free Cholesterol in Large VLDL	-0.011	(-0.023 ,0.002)	0.243
Triglycerides in Large VLDL	-0.012	(-0.025 ,0.000)	0.243
Concentration of Medium VLDL Particles	-0.014	(-0.027 ,0.000)	0.243
Total Lipids in Medium VLDL	-0.015	(-0.028 ,-0.001)	0.243
Phospholipids in Medium VLDL	-0.014	(-0.027 ,-0.001)	0.243
Cholesterol in Medium VLDL	-0.013	(-0.027 ,0.000)	0.243
Cholesteryl Esters in Medium VLDL	-0.012	(-0.025 ,0.001)	0.243
Free Cholesterol in Medium VLDL	-0.014	(-0.027 ,0.000)	0.243
Triglycerides in Medium VLDL	-0.013	(-0.026 ,0.000)	0.243
Concentration of Small VLDL Particles	-0.012	(-0.025 ,0.001)	0.243
Total Lipids in Small VLDL	-0.012	(-0.025 ,0.001)	0.243
Phospholipids in Small VLDL	-0.012	(-0.026 ,0.001)	0.243
Cholesterol in Small VLDL	-0.010	(-0.023 ,0.003)	0.267
Cholesteryl Esters in Small VLDL	-0.009	(-0.022 ,0.004)	0.303
Free Cholesterol in Small VLDL	-0.012	(-0.025 ,0.002)	0.243
Triglycerides in Small VLDL	-0.011	(-0.024 ,0.002)	0.243
Concentration of Very Small VLDL Particles	-0.010	(-0.023 ,0.003)	0.267
Total Lipids in Very Small VLDL	-0.010	(-0.023 ,0.003)	0.267

Phospholipids in Very Small VLDL	-0.010	(-0.023 ,0.003)	0.267
Cholesterol in Very Small VLDL	-0.009	(-0.023 ,0.004)	0.279
Cholesteryl Esters in Very Small VLDL	-0.009	(-0.022 ,0.004)	0.288
Free Cholesterol in Very Small VLDL	-0.010	(-0.023 ,0.004)	0.274
Triglycerides in Very Small VLDL	-0.007	(-0.020 ,0.006)	0.391
Concentration of IDL Particles	-0.006	(-0.020 ,0.007)	0.431
Total Lipids in IDL	-0.009	(-0.022 ,0.004)	0.303
Phospholipids in IDL	-0.011	(-0.024 ,0.002)	0.243
Cholesterol in IDL	-0.008	(-0.021 ,0.005)	0.368
Cholesteryl Esters in IDL	-0.007	(-0.020 ,0.006)	0.372
Free Cholesterol in IDL	-0.008	(-0.021 ,0.005)	0.340
Triglycerides in IDL	-0.006	(-0.019 ,0.007)	0.441
Concentration of Large LDL Particles	-0.014	(-0.028 ,-0.001)	0.243
Total Lipids in Large LDL	-0.011	(-0.024 ,0.003)	0.251
Phospholipids in Large LDL	-0.008	(-0.022 ,0.005)	0.340
Cholesterol in Large LDL	-0.011	(-0.025 ,0.002)	0.243
Cholesteryl Esters in Large LDL	-0.012	(-0.025 ,0.001)	0.243
Free Cholesterol in Large LDL	-0.010	(-0.023 ,0.003)	0.269
Triglycerides in Large LDL	-0.006	(-0.019 ,0.007)	0.456
Concentration of Medium LDL Particles	-0.011	(-0.025 ,0.002)	0.243
Total Lipids in Medium LDL	-0.010	(-0.024 ,0.003)	0.267
Phospholipids in Medium LDL	-0.009	(-0.023 ,0.004)	0.297
Cholesterol in Medium LDL	-0.010	(-0.024 ,0.003)	0.267
Cholesteryl Esters in Medium LDL	-0.011	(-0.024 ,0.003)	0.256
Free Cholesterol in Medium LDL	-0.009	(-0.022 ,0.005)	0.307
Triglycerides in Medium LDL	-0.008	(-0.021 ,0.005)	0.340

Concentration of Small LDL Particles	-0.015	(-0.029 ,-0.002)	0.243
Total Lipids in Small LDL	-0.013	(-0.027 ,0.000)	0.243
Phospholipids in Small LDL	-0.014	(-0.027 ,0.000)	0.243
Cholesterol in Small LDL	-0.012	(-0.026 ,0.001)	0.243
Cholesteryl Esters in Small LDL	-0.012	(-0.026 ,0.001)	0.243
Free Cholesterol in Small LDL	-0.011	(-0.024 ,0.003)	0.251
Triglycerides in Small LDL	-0.011	(-0.024 ,0.002)	0.243
Concentration of Very Large HDL Particles	-0.013	(-0.025 ,-0.002)	0.243
Total Lipids in Very Large HDL	-0.011	(-0.022 ,0.001)	0.243
Phospholipids in Very Large HDL	-0.010	(-0.021 ,0.001)	0.243
Cholesterol in Very Large HDL	-0.011	(-0.022 ,0.001)	0.243
Cholesteryl Esters in Very Large HDL	-0.010	(-0.021 ,0.001)	0.243
Free Cholesterol in Very Large HDL	-0.012	(-0.024 ,0.000)	0.243
Triglycerides in Very Large HDL	-0.018	(-0.031 ,-0.004)	0.240
Concentration of Large HDL Particles	-0.007	(-0.018 ,0.004)	0.340
Total Lipids in Large HDL	-0.005	(-0.016 ,0.006)	0.450
Phospholipids in Large HDL	-0.004	(-0.015 ,0.007)	0.593
Cholesterol in Large HDL	-0.006	(-0.017 ,0.005)	0.408
Cholesteryl Esters in Large HDL	-0.005	(-0.016 ,0.006)	0.420
Free Cholesterol in Large HDL	-0.006	(-0.017 ,0.005)	0.372
Triglycerides in Large HDL	-0.011	(-0.024 ,0.002)	0.243
Concentration of Medium HDL Particles	0.005	(-0.006 ,0.017)	0.446
Total Lipids in Medium HDL	0.007	(-0.005 ,0.019)	0.340
Phospholipids in Medium HDL	0.008	(-0.004 ,0.020)	0.295
Cholesterol in Medium HDL	0.007	(-0.005 ,0.019)	0.368
Cholesteryl Esters in Medium HDL	0.008	(-0.004 ,0.019)	0.306

Free Cholesterol in Medium HDL	0.003	(-0.009 ,0.015)	0.662
Triglycerides in Medium HDL	-0.002	(-0.015 ,0.011)	0.824
Concentration of Small HDL Particles	0.010	(-0.003 ,0.024)	0.256
Total Lipids in Small HDL	0.011	(-0.002 ,0.024)	0.243
Phospholipids in Small HDL	0.012	(-0.001 ,0.025)	0.243
Cholesterol in Small HDL	0.011	(-0.002 ,0.025)	0.243
Cholesteryl Esters in Small HDL	0.014	(0.000 ,0.027)	0.243
Free Cholesterol in Small HDL	0.003	(-0.010 ,0.016)	0.699
Triglycerides in Small HDL	-0.002	(-0.015 ,0.010)	0.732
Total Fatty Acids	-0.009	(-0.022 ,0.004)	0.301
Polyunsaturated Fatty Acids	-0.017	(-0.031 ,-0.004)	0.240
Monounsaturated Fatty Acids	-0.004	(-0.017 ,0.009)	0.635
Saturated Fatty Acids	-0.004	(-0.017 ,0.009)	0.604
Linoleic Acid	-0.023	(-0.036 ,-0.010)	0.168
Docosahexaenoic Acid	-0.014	(-0.027 ,-0.002)	0.243
Omega-3 Fatty Acids	-0.014	(-0.027 ,-0.001)	0.243
Omega-6 Fatty Acids	-0.016	(-0.029 ,-0.002)	0.243
Degree of Unsaturation	-0.006	(-0.019 ,0.006)	0.408
Total Cholines	-0.006	(-0.019 ,0.006)	0.417
Phosphatidylcholines	-0.008	(-0.020 ,0.005)	0.340
Sphingomyelins	-0.003	(-0.016 ,0.009)	0.659
Phosphoglycerides	-0.005	(-0.017 ,0.008)	0.520
Alanine	-0.016	(-0.030 ,-0.003)	0.243
Glutamine	0.000	(-0.013 ,0.014)	0.980
Glycine	-0.018	(-0.030 ,-0.005)	0.168
Histidine	-0.002	(-0.016 ,0.011)	0.787

Total Concentration of Branched-Chain Amino Acids (Leucine + Isoleucine + Valine)	-0.019	(-0.032 ,-0.006)	0.168
Isoleucine	-0.020	(-0.033 ,-0.007)	0.168
Leucine	-0.015	(-0.028 ,-0.002)	0.243
Valine	-0.020	(-0.032 ,-0.007)	0.168
Phenylalanine	-0.017	(-0.031 ,-0.004)	0.243
Tyrosine	-0.001	(-0.015 ,0.012)	0.864
Glucose	0.010	(-0.003 ,0.023)	0.267
Lactate	-0.005	(-0.019 ,0.008)	0.520
Pyruvate	-0.006	(-0.020 ,0.007)	0.441
Citrate	-0.012	(-0.025 ,0.001)	0.243
Acetone	-0.004	(-0.018 ,0.010)	0.634
Acetoacetate	0.002	(-0.012 ,0.015)	0.842
Acetate	-0.016	(-0.030 ,-0.003)	0.243
3-Hydroxybutyrate	0.004	(-0.010 ,0.017)	0.652
Glycoprotein Acetyls	0.007	(-0.005 ,0.020)	0.372
Creatinine	-0.004	(-0.016 ,0.007)	0.520
Albumin	-0.005	(-0.018 ,0.008)	0.532

COPD: chronic obstructive pulmonary disease; CI: confidential interval; FDR, false discovery rate; VLDL: Very-low-density lipoprotein; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Metabolites concentrations were standardized by z-score.

* indicates FDR-adjusted p values < 0.05.

Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.

Table S10. Associations of metabolites with COPD risk (n=23255)

Metabolites#	HR [a]	95% CI	P_{FDR}-value
Apolipoprotein A1	1.019	(0.925 ,1.121)	0.842
Apolipoprotein B	0.855	(0.786 ,0.929)	<0.001*
Total Concentration of Lipoprotein Particles	0.963	(0.879 ,1.056)	0.615
Concentration of VLDL Particles	0.897	(0.823 ,0.977)	0.035*
Concentration of LDL Particles	0.852	(0.783 ,0.926)	<0.001*
Concentration of HDL Particles	0.993	(0.905 ,1.089)	0.937
Average Diameter for VLDL Particles	0.951	(0.868 ,1.043)	0.458
Average Diameter for LDL Particles	0.971	(0.893 ,1.056)	0.669
Average Diameter for HDL Particles	1.013	(0.913 ,1.124)	0.933
Total Lipids in Lipoprotein Particles	0.888	(0.816 ,0.966)	0.018*
Total Lipids in VLDL	0.924	(0.847 ,1.008)	0.153
Total Lipids in LDL	0.854	(0.785 ,0.929)	<0.001*
Total Lipids in HDL	1.024	(0.928 ,1.131)	0.784
Total Cholesterol	0.857	(0.786 ,0.935)	<0.001*
VLDL Cholesterol	0.880	(0.808 ,0.958)	0.010*
LDL Cholesterol	0.849	(0.781 ,0.923)	<0.001*
HDL Cholesterol	1.002	(0.905 ,1.11)	0.974
Clinical LDL Cholesterol	0.847	(0.779 ,0.921)	<0.001*
Total Cholesterol Minus HDL-C	0.850	(0.781 ,0.925)	<0.001*
Remnant Cholesterol (Non-HDL, Non-LDL -Cholesterol)	0.856	(0.787 ,0.931)	<0.001*
Total Triglycerides	0.970	(0.891 ,1.056)	0.669
Triglycerides in VLDL	0.963	(0.884 ,1.05)	0.586
Triglycerides in LDL	0.995	(0.917 ,1.08)	0.951

Triglycerides in HDL	1.006	(0.927 ,1.092)	0.937
Total Phospholipids in Lipoprotein Particles	0.914	(0.836 ,0.998)	0.103
Phospholipids in VLDL	0.912	(0.837 ,0.994)	0.084
Phospholipids in LDL	0.854	(0.785 ,0.928)	<0.001*
Phospholipids in HDL	1.044	(0.948 ,1.15)	0.584
Total Esterified Cholesterol	0.858	(0.787 ,0.937)	0.004*
Cholesteryl Esters in VLDL	0.871	(0.8 ,0.948)	0.004*
Cholesteryl Esters in LDL	0.853	(0.784 ,0.928)	<0.001*
Cholesteryl Esters in HDL	1.006	(0.909 ,1.113)	0.951
Total Free Cholesterol	0.856	(0.786 ,0.932)	<0.001*
Free Cholesterol in VLDL	0.897	(0.823 ,0.977)	0.035*
Free Cholesterol in LDL	0.843	(0.775 ,0.916)	<0.001*
Free Cholesterol in HDL	0.990	(0.894 ,1.097)	0.933
Concentration of Chylomicrons and Extremely Large VLDL Particles	1.002	(0.922 ,1.089)	0.974
Total Lipids in Chylomicrons and Extremely Large VLDL	0.997	(0.917 ,1.083)	0.973
Phospholipids in Chylomicrons and Extremely Large VLDL	1.007	(0.926 ,1.095)	0.937
Cholesterol in Chylomicrons and Extremely Large VLDL	0.985	(0.905 ,1.073)	0.860
Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	0.970	(0.89 ,1.057)	0.669
Free Cholesterol in Chylomicrons and Extremely Large VLDL	1.003	(0.922 ,1.091)	0.973
Triglycerides in Chylomicrons and Extremely Large VLDL	0.998	(0.919 ,1.084)	0.974
Concentration of Very Large VLDL Particles	0.956	(0.876 ,1.044)	0.498
Total Lipids in Very Large VLDL	0.957	(0.877 ,1.044)	0.501
Phospholipids in Very Large VLDL	0.950	(0.87 ,1.037)	0.422
Cholesterol in Very Large VLDL	0.915	(0.837 ,0.999)	0.104
Cholesteryl Esters in Very Large VLDL	0.893	(0.818 ,0.975)	0.033*
Free Cholesterol in Very Large VLDL	0.941	(0.862 ,1.027)	0.309

Triglycerides in Very Large VLDL	0.976	(0.895 ,1.064)	0.744
Concentration of Large VLDL Particles	0.934	(0.855 ,1.02)	0.243
Total Lipids in Large VLDL	0.929	(0.851 ,1.015)	0.196
Phospholipids in Large VLDL	0.939	(0.859 ,1.026)	0.301
Cholesterol in Large VLDL	0.905	(0.829 ,0.989)	0.067
Cholesteryl Esters in Large VLDL	0.888	(0.813 ,0.969)	0.023*
Free Cholesterol in Large VLDL	0.926	(0.847 ,1.012)	0.176
Triglycerides in Large VLDL	0.940	(0.862 ,1.026)	0.305
Concentration of Medium VLDL Particles	0.864	(0.793 ,0.941)	0.004*
Total Lipids in Medium VLDL	0.876	(0.804 ,0.955)	0.010*
Phospholipids in Medium VLDL	0.865	(0.794 ,0.942)	0.004*
Cholesterol in Medium VLDL	0.847	(0.779 ,0.922)	<0.001*
Cholesteryl Esters in Medium VLDL	0.846	(0.778 ,0.92)	<0.001*
Free Cholesterol in Medium VLDL	0.857	(0.787 ,0.933)	<0.001*
Triglycerides in Medium VLDL	0.925	(0.848 ,1.009)	0.160
Concentration of Small VLDL Particles	0.908	(0.833 ,0.991)	0.071
Total Lipids in Small VLDL	0.907	(0.832 ,0.989)	0.066
Phospholipids in Small VLDL	0.878	(0.807 ,0.957)	0.010*
Cholesterol in Small VLDL	0.876	(0.804 ,0.954)	0.007*
Cholesteryl Esters in Small VLDL	0.885	(0.812 ,0.964)	0.015*
Free Cholesterol in Small VLDL	0.863	(0.793 ,0.939)	0.004*
Triglycerides in Small VLDL	0.973	(0.894 ,1.06)	0.698
Concentration of Very Small VLDL Particles	0.907	(0.833 ,0.987)	0.060
Total Lipids in Very Small VLDL	0.914	(0.839 ,0.995)	0.086
Phospholipids in Very Small VLDL	0.926	(0.85 ,1.008)	0.153
Cholesterol in Very Small VLDL	0.882	(0.809 ,0.961)	0.012*

Cholesteryl Esters in Very Small VLDL	0.875	(0.802 ,0.954)	0.007*
Free Cholesterol in Very Small VLDL	0.903	(0.829 ,0.984)	0.050*
Triglycerides in Very Small VLDL	1.010	(0.93 ,1.097)	0.933
Concentration of IDL Particles	0.863	(0.793 ,0.938)	<0.001*
Total Lipids in IDL	0.856	(0.785 ,0.933)	<0.001*
Phospholipids in IDL	0.863	(0.792 ,0.941)	0.004*
Cholesterol in IDL	0.847	(0.777 ,0.923)	<0.001*
Cholesteryl Esters in IDL	0.845	(0.775 ,0.921)	<0.001*
Free Cholesterol in IDL	0.856	(0.786 ,0.933)	<0.001*
Triglycerides in IDL	1.010	(0.931 ,1.097)	0.933
Concentration of Large LDL Particles	0.853	(0.784 ,0.927)	<0.001*
Total Lipids in Large LDL	0.851	(0.782 ,0.926)	<0.001*
Phospholipids in Large LDL	0.852	(0.784 ,0.927)	<0.001*
Cholesterol in Large LDL	0.845	(0.776 ,0.92)	<0.001*
Cholesteryl Esters in Large LDL	0.848	(0.779 ,0.923)	<0.001*
Free Cholesterol in Large LDL	0.841	(0.772 ,0.915)	<0.001*
Triglycerides in Large LDL	1.000	(0.921 ,1.086)	0.998
Concentration of Medium LDL Particles	0.860	(0.79 ,0.935)	<0.001*
Total Lipids in Medium LDL	0.863	(0.794 ,0.938)	<0.001*
Phospholipids in Medium LDL	0.857	(0.788 ,0.932)	<0.001*
Cholesterol in Medium LDL	0.860	(0.791 ,0.935)	<0.001*
Cholesteryl Esters in Medium LDL	0.867	(0.797 ,0.943)	0.004*
Free Cholesterol in Medium LDL	0.848	(0.781 ,0.921)	<0.001*
Triglycerides in Medium LDL	0.991	(0.913 ,1.076)	0.933
Concentration of Small LDL Particles	0.856	(0.787 ,0.93)	<0.001*
Total Lipids in Small LDL	0.859	(0.791 ,0.933)	<0.001*

Phospholipids in Small LDL	0.865	(0.797 ,0.939)	<0.001*
Cholesterol in Small LDL	0.852	(0.785 ,0.926)	<0.001*
Cholesteryl Esters in Small LDL	0.859	(0.79 ,0.934)	<0.001*
Free Cholesterol in Small LDL	0.850	(0.784 ,0.922)	<0.001*
Triglycerides in Small LDL	0.983	(0.905 ,1.067)	0.816
Concentration of Very Large HDL Particles	0.934	(0.841 ,1.038)	0.353
Total Lipids in Very Large HDL	0.959	(0.863 ,1.065)	0.621
Phospholipids in Very Large HDL	0.976	(0.879 ,1.083)	0.784
Cholesterol in Very Large HDL	0.942	(0.848 ,1.047)	0.434
Cholesteryl Esters in Very Large HDL	0.944	(0.849 ,1.05)	0.463
Free Cholesterol in Very Large HDL	0.937	(0.847 ,1.035)	0.348
Triglycerides in Very Large HDL	0.966	(0.89 ,1.049)	0.601
Concentration of Large HDL Particles	0.989	(0.89 ,1.1)	0.933
Total Lipids in Large HDL	1.007	(0.907 ,1.119)	0.944
Phospholipids in Large HDL	1.027	(0.926 ,1.139)	0.773
Cholesterol in Large HDL	0.990	(0.89 ,1.101)	0.933
Cholesteryl Esters in Large HDL	0.990	(0.89 ,1.101)	0.933
Free Cholesterol in Large HDL	0.992	(0.892 ,1.102)	0.937
Triglycerides in Large HDL	0.990	(0.909 ,1.078)	0.933
Concentration of Medium HDL Particles	1.042	(0.949 ,1.145)	0.585
Total Lipids in Medium HDL	1.058	(0.966 ,1.16)	0.376
Phospholipids in Medium HDL	1.069	(0.978 ,1.17)	0.262
Cholesterol in Medium HDL	1.047	(0.952 ,1.151)	0.524
Cholesteryl Esters in Medium HDL	1.051	(0.956 ,1.155)	0.473
Free Cholesterol in Medium HDL	1.029	(0.934 ,1.133)	0.731
Triglycerides in Medium HDL	1.020	(0.94 ,1.106)	0.784

Concentration of Small HDL Particles	0.973	(0.893 ,1.059)	0.698
Total Lipids in Small HDL	1.010	(0.928 ,1.099)	0.933
Phospholipids in Small HDL	1.029	(0.946 ,1.12)	0.677
Cholesterol in Small HDL	0.979	(0.899 ,1.065)	0.774
Cholesteryl Esters in Small HDL	0.982	(0.903 ,1.069)	0.817
Free Cholesterol in Small HDL	0.970	(0.89 ,1.057)	0.669
Triglycerides in Small HDL	1.009	(0.925 ,1.1)	0.933
Total Fatty Acids	0.954	(0.877 ,1.037)	0.434
Polyunsaturated Fatty Acids	0.881	(0.808 ,0.962)	0.015*
Monounsaturated Fatty Acids	1.008	(0.93 ,1.093)	0.933
Saturated Fatty Acids	0.974	(0.897 ,1.058)	0.698
Linoleic Acid	0.869	(0.798 ,0.947)	0.004*
Docosahexaenoic Acid	0.882	(0.803 ,0.969)	0.026*
Omega-3 Fatty Acids	0.899	(0.821 ,0.984)	0.054
Omega-6 Fatty Acids	0.894	(0.82 ,0.974)	0.031*
Degree of Unsaturation	0.903	(0.827 ,0.986)	0.060
Total Cholines	0.925	(0.845 ,1.014)	0.181
Phosphatidylcholines	0.924	(0.843 ,1.013)	0.176
Sphingomyelins	0.922	(0.842 ,1.01)	0.160
Phosphoglycerides	0.942	(0.861 ,1.031)	0.340
Alanine	0.870	(0.799 ,0.947)	0.004*
Glutamine	0.922	(0.849 ,1.002)	0.123
Glycine	0.909	(0.824 ,1.002)	0.120
Histidine	0.926	(0.849 ,1.01)	0.164
Total Concentration of Branched-Chain Amino Acids (Leucine + Isoleucine + Valine)	0.976	(0.895 ,1.065)	0.747

Isoleucine	1.002	(0.922 ,1.09)	0.974
Leucine	0.967	(0.886 ,1.056)	0.642
Valine	0.973	(0.891 ,1.062)	0.698
Phenylalanine	1.033	(0.954 ,1.119)	0.611
Tyrosine	1.088	(1.005 ,1.178)	0.086
Glucose	0.981	(0.908 ,1.06)	0.777
Lactate	1.073	(0.992 ,1.161)	0.160
Pyruvate	1.034	(0.952 ,1.124)	0.611
Citrate	0.905	(0.829 ,0.988)	0.064
Acetone	1.044	(0.977 ,1.115)	0.353
Acetoacetate	1.069	(0.997 ,1.146)	0.125
Acetate	0.915	(0.801 ,1.044)	0.327
3-Hydroxybutyrate	1.042	(0.968 ,1.123)	0.441
Glycoprotein Acetyls	1.229	(1.133 ,1.335)	<0.001*
Creatinine	1.025	(0.95 ,1.106)	0.698
Albumin	0.823	(0.757 ,0.894)	<0.001*

COPD: chronic obstructive pulmonary disease; HR: hazard ratio; CI: confidential interval; FDR, false discovery rate; VLDL: Very-low-density lipoprotein; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Metabolites concentrations were standardized by z-score.

* indicates FDR-adjusted p values < 0.05.

Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.

Table S11. Associations of total lifetime occupational hazard exposure with metabolites after removing baseline CVD cases (n=19718)

Metabolites#	beta [a]	95% CI	P_{FDR}-value
Apolipoprotein A1	0.001	(-0.011 ,0.014)	0.869
Apolipoprotein B	-0.010	(-0.025 ,0.004)	0.422
Total Concentration of Lipoprotein Particles	0.001	(-0.012 ,0.015)	0.875
Concentration of VLDL Particles	-0.012	(-0.026 ,0.002)	0.422
Concentration of LDL Particles	-0.011	(-0.026 ,0.003)	0.422
Concentration of HDL Particles	0.003	(-0.010 ,0.016)	0.728
Average Diameter for VLDL Particles	-0.009	(-0.022 ,0.003)	0.422
Average Diameter for LDL Particles	0.002	(-0.012 ,0.016)	0.835
Average Diameter for HDL Particles	-0.006	(-0.018 ,0.006)	0.508
Total Lipids in Lipoprotein Particles	-0.010	(-0.025 ,0.004)	0.422
Total Lipids in VLDL	-0.012	(-0.026 ,0.002)	0.422
Total Lipids in LDL	-0.008	(-0.023 ,0.006)	0.453
Total Lipids in HDL	0.001	(-0.012 ,0.013)	0.940
Total Cholesterol	-0.008	(-0.022 ,0.006)	0.460
VLDL Cholesterol	-0.011	(-0.026 ,0.003)	0.422
LDL Cholesterol	-0.009	(-0.023 ,0.006)	0.449
HDL Cholesterol	0.001	(-0.011 ,0.013)	0.875
Clinical LDL Cholesterol	-0.007	(-0.021 ,0.008)	0.508
Total Cholesterol Minus HDL-C	-0.009	(-0.024 ,0.005)	0.436
Remnant Cholesterol (Non-HDL, Non-LDL -Cholesterol)	-0.010	(-0.024 ,0.005)	0.422
Total Triglycerides	-0.011	(-0.025 ,0.003)	0.422
Triglycerides in VLDL	-0.011	(-0.025 ,0.002)	0.422
Triglycerides in LDL	-0.008	(-0.022 ,0.007)	0.486

Triglycerides in HDL	-0.010	(-0.024 ,0.005)	0.422
Total Phospholipids in Lipoprotein Particles	-0.007	(-0.021 ,0.007)	0.508
Phospholipids in VLDL	-0.012	(-0.026 ,0.002)	0.422
Phospholipids in LDL	-0.007	(-0.021 ,0.008)	0.508
Phospholipids in HDL	0.001	(-0.011 ,0.014)	0.869
Total Esterified Cholesterol	-0.007	(-0.021 ,0.007)	0.508
Cholesteryl Esters in VLDL	-0.011	(-0.025 ,0.003)	0.422
Cholesteryl Esters in LDL	-0.009	(-0.024 ,0.005)	0.433
Cholesteryl Esters in HDL	0.002	(-0.010 ,0.014)	0.818
Total Free Cholesterol	-0.010	(-0.024 ,0.004)	0.422
Free Cholesterol in VLDL	-0.012	(-0.026 ,0.002)	0.422
Free Cholesterol in LDL	-0.006	(-0.021 ,0.008)	0.527
Free Cholesterol in HDL	-0.002	(-0.014 ,0.010)	0.835
Concentration of Chylomicrons and Extremely Large VLDL Particles	-0.010	(-0.023 ,0.004)	0.422
Total Lipids in Chylomicrons and Extremely Large VLDL	-0.010	(-0.023 ,0.004)	0.422
Phospholipids in Chylomicrons and Extremely Large VLDL	-0.010	(-0.023 ,0.004)	0.422
Cholesterol in Chylomicrons and Extremely Large VLDL	-0.009	(-0.023 ,0.004)	0.422
Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	-0.009	(-0.022 ,0.005)	0.422
Free Cholesterol in Chylomicrons and Extremely Large VLDL	-0.010	(-0.023 ,0.004)	0.422
Triglycerides in Chylomicrons and Extremely Large VLDL	-0.010	(-0.023 ,0.004)	0.422
Concentration of Very Large VLDL Particles	-0.010	(-0.024 ,0.003)	0.422
Total Lipids in Very Large VLDL	-0.010	(-0.024 ,0.003)	0.422
Phospholipids in Very Large VLDL	-0.011	(-0.024 ,0.003)	0.422
Cholesterol in Very Large VLDL	-0.011	(-0.024 ,0.003)	0.422
Cholesteryl Esters in Very Large VLDL	-0.010	(-0.024 ,0.003)	0.422

Free Cholesterol in Very Large VLDL	-0.011	(-0.025 ,0.002)	0.422
Triglycerides in Very Large VLDL	-0.010	(-0.023 ,0.004)	0.422
Concentration of Large VLDL Particles	-0.011	(-0.024 ,0.002)	0.422
Total Lipids in Large VLDL	-0.011	(-0.025 ,0.002)	0.422
Phospholipids in Large VLDL	-0.011	(-0.024 ,0.003)	0.422
Cholesterol in Large VLDL	-0.011	(-0.025 ,0.002)	0.422
Cholesteryl Esters in Large VLDL	-0.012	(-0.025 ,0.002)	0.422
Free Cholesterol in Large VLDL	-0.011	(-0.024 ,0.003)	0.422
Triglycerides in Large VLDL	-0.011	(-0.025 ,0.002)	0.422
Concentration of Medium VLDL Particles	-0.012	(-0.026 ,0.002)	0.422
Total Lipids in Medium VLDL	-0.013	(-0.027 ,0.001)	0.422
Phospholipids in Medium VLDL	-0.012	(-0.027 ,0.002)	0.422
Cholesterol in Medium VLDL	-0.011	(-0.025 ,0.004)	0.422
Cholesteryl Esters in Medium VLDL	-0.010	(-0.024 ,0.005)	0.422
Free Cholesterol in Medium VLDL	-0.012	(-0.026 ,0.003)	0.422
Triglycerides in Medium VLDL	-0.012	(-0.026 ,0.002)	0.422
Concentration of Small VLDL Particles	-0.011	(-0.025 ,0.003)	0.422
Total Lipids in Small VLDL	-0.011	(-0.025 ,0.003)	0.422
Phospholipids in Small VLDL	-0.011	(-0.025 ,0.003)	0.422
Cholesterol in Small VLDL	-0.009	(-0.023 ,0.005)	0.448
Cholesteryl Esters in Small VLDL	-0.008	(-0.023 ,0.006)	0.449
Free Cholesterol in Small VLDL	-0.010	(-0.024 ,0.005)	0.422
Triglycerides in Small VLDL	-0.012	(-0.025 ,0.002)	0.422
Concentration of Very Small VLDL Particles	-0.010	(-0.024 ,0.005)	0.422
Total Lipids in Very Small VLDL	-0.010	(-0.024 ,0.005)	0.422
Phospholipids in Very Small VLDL	-0.010	(-0.024 ,0.005)	0.422

Cholesterol in Very Small VLDL	-0.009	(-0.023 ,0.006)	0.449
Cholesteryl Esters in Very Small VLDL	-0.008	(-0.023 ,0.006)	0.449
Free Cholesterol in Very Small VLDL	-0.009	(-0.023 ,0.006)	0.449
Triglycerides in Very Small VLDL	-0.008	(-0.023 ,0.006)	0.449
Concentration of IDL Particles	-0.004	(-0.019 ,0.010)	0.626
Total Lipids in IDL	-0.007	(-0.021 ,0.007)	0.508
Phospholipids in IDL	-0.009	(-0.023 ,0.005)	0.422
Cholesterol in IDL	-0.005	(-0.019 ,0.009)	0.568
Cholesteryl Esters in IDL	-0.005	(-0.019 ,0.009)	0.570
Free Cholesterol in IDL	-0.006	(-0.020 ,0.008)	0.546
Triglycerides in IDL	-0.007	(-0.021 ,0.007)	0.508
Concentration of Large LDL Particles	-0.011	(-0.025 ,0.004)	0.422
Total Lipids in Large LDL	-0.008	(-0.022 ,0.007)	0.479
Phospholipids in Large LDL	-0.006	(-0.021 ,0.008)	0.527
Cholesterol in Large LDL	-0.008	(-0.023 ,0.006)	0.456
Cholesteryl Esters in Large LDL	-0.009	(-0.023 ,0.006)	0.449
Free Cholesterol in Large LDL	-0.006	(-0.021 ,0.008)	0.508
Triglycerides in Large LDL	-0.006	(-0.020 ,0.008)	0.527
Concentration of Medium LDL Particles	-0.011	(-0.025 ,0.004)	0.422
Total Lipids in Medium LDL	-0.008	(-0.023 ,0.006)	0.453
Phospholipids in Medium LDL	-0.006	(-0.021 ,0.008)	0.508
Cholesterol in Medium LDL	-0.009	(-0.023 ,0.006)	0.449
Cholesteryl Esters in Medium LDL	-0.010	(-0.024 ,0.005)	0.422
Free Cholesterol in Medium LDL	-0.005	(-0.020 ,0.009)	0.570
Triglycerides in Medium LDL	-0.008	(-0.022 ,0.006)	0.449
Concentration of Small LDL Particles	-0.013	(-0.027 ,0.002)	0.422

Total Lipids in Small LDL	-0.010	(-0.024 ,0.005)	0.422
Phospholipids in Small LDL	-0.008	(-0.023 ,0.006)	0.453
Cholesterol in Small LDL	-0.010	(-0.024 ,0.005)	0.422
Cholesteryl Esters in Small LDL	-0.011	(-0.025 ,0.004)	0.422
Free Cholesterol in Small LDL	-0.006	(-0.021 ,0.009)	0.537
Triglycerides in Small LDL	-0.012	(-0.026 ,0.002)	0.422
Concentration of Very Large HDL Particles	-0.008	(-0.021 ,0.004)	0.422
Total Lipids in Very Large HDL	-0.006	(-0.019 ,0.006)	0.508
Phospholipids in Very Large HDL	-0.006	(-0.018 ,0.007)	0.508
Cholesterol in Very Large HDL	-0.006	(-0.018 ,0.007)	0.508
Cholesteryl Esters in Very Large HDL	-0.005	(-0.018 ,0.007)	0.508
Free Cholesterol in Very Large HDL	-0.006	(-0.019 ,0.007)	0.508
Triglycerides in Very Large HDL	-0.018	(-0.032 ,-0.003)	0.422
Concentration of Large HDL Particles	-0.004	(-0.016 ,0.007)	0.570
Total Lipids in Large HDL	-0.003	(-0.015 ,0.009)	0.703
Phospholipids in Large HDL	-0.002	(-0.014 ,0.010)	0.789
Cholesterol in Large HDL	-0.003	(-0.015 ,0.009)	0.705
Cholesteryl Esters in Large HDL	-0.003	(-0.015 ,0.009)	0.724
Free Cholesterol in Large HDL	-0.004	(-0.016 ,0.008)	0.626
Triglycerides in Large HDL	-0.013	(-0.027 ,0.001)	0.422
Concentration of Medium HDL Particles	0.003	(-0.010 ,0.016)	0.724
Total Lipids in Medium HDL	0.004	(-0.009 ,0.017)	0.610
Phospholipids in Medium HDL	0.005	(-0.009 ,0.018)	0.592
Cholesterol in Medium HDL	0.005	(-0.008 ,0.018)	0.547
Cholesteryl Esters in Medium HDL	0.006	(-0.007 ,0.019)	0.508
Free Cholesterol in Medium HDL	0.002	(-0.011 ,0.014)	0.828

Triglycerides in Medium HDL	-0.007	(-0.021 ,0.007)	0.508
Concentration of Small HDL Particles	0.007	(-0.007 ,0.021)	0.508
Total Lipids in Small HDL	0.006	(-0.008 ,0.021)	0.508
Phospholipids in Small HDL	0.007	(-0.007 ,0.021)	0.508
Cholesterol in Small HDL	0.008	(-0.006 ,0.023)	0.453
Cholesteryl Esters in Small HDL	0.010	(-0.004 ,0.025)	0.422
Free Cholesterol in Small HDL	0.001	(-0.013 ,0.015)	0.875
Triglycerides in Small HDL	-0.006	(-0.020 ,0.007)	0.508
Total Fatty Acids	-0.010	(-0.025 ,0.004)	0.422
Polyunsaturated Fatty Acids	-0.018	(-0.032 ,-0.004)	0.422
Monounsaturated Fatty Acids	-0.006	(-0.020 ,0.008)	0.508
Saturated Fatty Acids	-0.005	(-0.019 ,0.009)	0.572
Linoleic Acid	-0.023	(-0.037 ,-0.008)	0.336
Docosahexaenoic Acid	-0.013	(-0.026 ,0.001)	0.422
Omega-3 Fatty Acids	-0.016	(-0.030 ,-0.002)	0.422
Omega-6 Fatty Acids	-0.016	(-0.030 ,-0.002)	0.422
Degree of Unsaturation	-0.008	(-0.021 ,0.005)	0.449
Total Cholines	-0.007	(-0.020 ,0.007)	0.508
Phosphatidylcholines	-0.008	(-0.022 ,0.005)	0.445
Sphingomyelins	0.000	(-0.014 ,0.013)	0.970
Phosphoglycerides	-0.006	(-0.020 ,0.007)	0.508
Alanine	-0.014	(-0.029 ,0.000)	0.422
Glutamine	0.002	(-0.012 ,0.017)	0.831
Glycine	-0.013	(-0.026 ,0.001)	0.422
Histidine	-0.007	(-0.021 ,0.008)	0.508

Total Concentration of Branched-Chain Amino Acids (Leucine + Isoleucine + Valine)	-0.015	(-0.028 ,-0.001)	0.422
Isoleucine	-0.017	(-0.031 ,-0.003)	0.422
Leucine	-0.011	(-0.025 ,0.003)	0.422
Valine	-0.015	(-0.028 ,-0.001)	0.422
Phenylalanine	-0.013	(-0.028 ,0.002)	0.422
Tyrosine	0.000	(-0.015 ,0.014)	0.970
Glucose	0.014	(-0.001 ,0.028)	0.422
Lactate	-0.008	(-0.022 ,0.007)	0.491
Pyruvate	-0.004	(-0.019 ,0.010)	0.664
Citrate	-0.011	(-0.025 ,0.003)	0.422
Acetone	-0.003	(-0.017 ,0.012)	0.765
Acetoacetate	0.002	(-0.013 ,0.016)	0.862
Acetate	-0.020	(-0.035 ,-0.005)	0.422
3-Hydroxybutyrate	0.005	(-0.010 ,0.019)	0.607
Glycoprotein Acetyls	0.005	(-0.008 ,0.019)	0.568
Creatinine	-0.005	(-0.017 ,0.007)	0.550
Albumin	-0.005	(-0.020 ,0.009)	0.570

CVD: cardiovascular disease; CI: confidential interval; FDR, false discovery rate; VLDL: Very-low-density lipoprotein; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Metabolites concentrations were standardized by z-score.

* indicates FDR-adjusted p values < 0.05.

Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.

Table S12. Associations of metabolites with CVD risk (n=19718)

Metabolites#	HR [a]	95% CI	P_{FDR}-value
Apolipoprotein A1	0.997	(0.962 ,1.035)	0.936
Apolipoprotein B	0.960	(0.931 ,0.991)	0.057
Total Concentration of Lipoprotein Particles	0.978	(0.945 ,1.013)	0.375
Concentration of VLDL Particles	0.971	(0.941 ,1.003)	0.186
Concentration of LDL Particles	0.960	(0.931 ,0.991)	0.057
Concentration of HDL Particles	0.985	(0.951 ,1.02)	0.560
Average Diameter for VLDL Particles	0.969	(0.936 ,1.004)	0.210
Average Diameter for LDL Particles	0.982	(0.951 ,1.013)	0.423
Average Diameter for HDL Particles	1.021	(0.982 ,1.061)	0.479
Total Lipids in Lipoprotein Particles	0.967	(0.936 ,0.998)	0.100
Total Lipids in VLDL	0.973	(0.941 ,1.005)	0.226
Total Lipids in LDL	0.961	(0.932 ,0.992)	0.057
Total Lipids in HDL	1.002	(0.965 ,1.041)	0.936
Total Cholesterol	0.960	(0.93 ,0.992)	0.057
VLDL Cholesterol	0.962	(0.932 ,0.994)	0.066
LDL Cholesterol	0.960	(0.931 ,0.99)	0.057
HDL Cholesterol	0.998	(0.96 ,1.037)	0.936
Clinical LDL Cholesterol	0.958	(0.928 ,0.988)	0.057
Total Cholesterol Minus HDL-C	0.959	(0.93 ,0.99)	0.057
Remnant Cholesterol (Non-HDL, Non-LDL -Cholesterol)	0.960	(0.93 ,0.99)	0.057
Total Triglycerides	0.986	(0.955 ,1.019)	0.571
Triglycerides in VLDL	0.983	(0.951 ,1.016)	0.477
Triglycerides in LDL	1.001	(0.97 ,1.033)	0.962

Triglycerides in HDL	1.003	(0.972 ,1.035)	0.925
Total Phospholipids in Lipoprotein Particles	0.973	(0.941 ,1.005)	0.227
Phospholipids in VLDL	0.970	(0.939 ,1.002)	0.178
Phospholipids in LDL	0.960	(0.93 ,0.99)	0.057
Phospholipids in HDL	1.007	(0.97 ,1.045)	0.866
Total Esterified Cholesterol	0.961	(0.93 ,0.992)	0.057
Cholesteryl Esters in VLDL	0.961	(0.931 ,0.992)	0.057
Cholesteryl Esters in LDL	0.961	(0.931 ,0.991)	0.057
Cholesteryl Esters in HDL	0.996	(0.959 ,1.035)	0.925
Total Free Cholesterol	0.960	(0.93 ,0.991)	0.057
Free Cholesterol in VLDL	0.966	(0.935 ,0.997)	0.100
Free Cholesterol in LDL	0.959	(0.93 ,0.989)	0.057
Free Cholesterol in HDL	1.002	(0.965 ,1.042)	0.936
Concentration of Chylomicrons and Extremely Large VLDL Particles	0.985	(0.953 ,1.018)	0.540
Total Lipids in Chylomicrons and Extremely Large VLDL	0.984	(0.952 ,1.017)	0.500
Phospholipids in Chylomicrons and Extremely Large VLDL	0.984	(0.952 ,1.017)	0.496
Cholesterol in Chylomicrons and Extremely Large VLDL	0.976	(0.944 ,1.009)	0.284
Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	0.972	(0.941 ,1.005)	0.227
Free Cholesterol in Chylomicrons and Extremely Large VLDL	0.980	(0.948 ,1.013)	0.405
Triglycerides in Chylomicrons and Extremely Large VLDL	0.987	(0.955 ,1.02)	0.588
Concentration of Very Large VLDL Particles	0.980	(0.948 ,1.014)	0.414
Total Lipids in Very Large VLDL	0.979	(0.947 ,1.013)	0.382
Phospholipids in Very Large VLDL	0.977	(0.945 ,1.011)	0.329
Cholesterol in Very Large VLDL	0.967	(0.935 ,1)	0.133
Cholesteryl Esters in Very Large VLDL	0.961	(0.93 ,0.994)	0.069
Free Cholesterol in Very Large VLDL	0.974	(0.942 ,1.007)	0.258

Triglycerides in Very Large VLDL	0.985	(0.953 ,1.018)	0.540
Concentration of Large VLDL Particles	0.976	(0.944 ,1.009)	0.294
Total Lipids in Large VLDL	0.975	(0.943 ,1.008)	0.265
Phospholipids in Large VLDL	0.977	(0.944 ,1.01)	0.319
Cholesterol in Large VLDL	0.968	(0.937 ,1.001)	0.152
Cholesteryl Esters in Large VLDL	0.965	(0.934 ,0.997)	0.096
Free Cholesterol in Large VLDL	0.973	(0.94 ,1.006)	0.228
Triglycerides in Large VLDL	0.978	(0.946 ,1.011)	0.356
Concentration of Medium VLDL Particles	0.960	(0.93 ,0.991)	0.057
Total Lipids in Medium VLDL	0.963	(0.932 ,0.994)	0.069
Phospholipids in Medium VLDL	0.960	(0.93 ,0.991)	0.057
Cholesterol in Medium VLDL	0.956	(0.926 ,0.986)	0.057
Cholesteryl Esters in Medium VLDL	0.955	(0.926 ,0.985)	0.057
Free Cholesterol in Medium VLDL	0.959	(0.929 ,0.989)	0.057
Triglycerides in Medium VLDL	0.975	(0.944 ,1.008)	0.268
Concentration of Small VLDL Particles	0.975	(0.944 ,1.007)	0.249
Total Lipids in Small VLDL	0.973	(0.942 ,1.005)	0.225
Phospholipids in Small VLDL	0.966	(0.936 ,0.997)	0.096
Cholesterol in Small VLDL	0.965	(0.935 ,0.996)	0.094
Cholesteryl Esters in Small VLDL	0.968	(0.937 ,0.999)	0.127
Free Cholesterol in Small VLDL	0.961	(0.932 ,0.992)	0.057
Triglycerides in Small VLDL	0.989	(0.958 ,1.022)	0.693
Concentration of Very Small VLDL Particles	0.980	(0.95 ,1.012)	0.382
Total Lipids in Very Small VLDL	0.983	(0.952 ,1.015)	0.468
Phospholipids in Very Small VLDL	0.987	(0.956 ,1.018)	0.571
Cholesterol in Very Small VLDL	0.974	(0.943 ,1.005)	0.227

Cholesteryl Esters in Very Small VLDL	0.971	(0.941 ,1.003)	0.191
Free Cholesterol in Very Small VLDL	0.981	(0.95 ,1.012)	0.390
Triglycerides in Very Small VLDL	1.006	(0.975 ,1.038)	0.857
Concentration of IDL Particles	0.961	(0.932 ,0.991)	0.057
Total Lipids in IDL	0.965	(0.935 ,0.996)	0.092
Phospholipids in IDL	0.967	(0.937 ,0.999)	0.115
Cholesterol in IDL	0.962	(0.931 ,0.993)	0.063
Cholesteryl Esters in IDL	0.961	(0.93 ,0.992)	0.057
Free Cholesterol in IDL	0.965	(0.935 ,0.997)	0.095
Triglycerides in IDL	1.008	(0.977 ,1.04)	0.779
Concentration of Large LDL Particles	0.960	(0.931 ,0.99)	0.057
Total Lipids in Large LDL	0.962	(0.932 ,0.993)	0.057
Phospholipids in Large LDL	0.959	(0.929 ,0.989)	0.057
Cholesterol in Large LDL	0.961	(0.931 ,0.991)	0.057
Cholesteryl Esters in Large LDL	0.962	(0.932 ,0.992)	0.057
Free Cholesterol in Large LDL	0.959	(0.93 ,0.99)	0.057
Triglycerides in Large LDL	1.005	(0.974 ,1.037)	0.875
Concentration of Medium LDL Particles	0.964	(0.934 ,0.994)	0.071
Total Lipids in Medium LDL	0.961	(0.931 ,0.991)	0.057
Phospholipids in Medium LDL	0.959	(0.93 ,0.99)	0.057
Cholesterol in Medium LDL	0.960	(0.93 ,0.99)	0.057
Cholesteryl Esters in Medium LDL	0.962	(0.932 ,0.992)	0.057
Free Cholesterol in Medium LDL	0.958	(0.929 ,0.988)	0.057
Triglycerides in Medium LDL	0.997	(0.966 ,1.029)	0.934
Concentration of Small LDL Particles	0.963	(0.934 ,0.994)	0.068
Total Lipids in Small LDL	0.961	(0.932 ,0.991)	0.057

Phospholipids in Small LDL	0.968	(0.938 ,0.998)	0.100
Cholesterol in Small LDL	0.958	(0.928 ,0.988)	0.057
Cholesteryl Esters in Small LDL	0.958	(0.928 ,0.988)	0.057
Free Cholesterol in Small LDL	0.961	(0.932 ,0.991)	0.057
Triglycerides in Small LDL	0.989	(0.958 ,1.022)	0.689
Concentration of Very Large HDL Particles	1.004	(0.967 ,1.043)	0.920
Total Lipids in Very Large HDL	1.010	(0.972 ,1.05)	0.751
Phospholipids in Very Large HDL	1.015	(0.977 ,1.054)	0.615
Cholesterol in Very Large HDL	1.005	(0.967 ,1.044)	0.913
Cholesteryl Esters in Very Large HDL	1.004	(0.966 ,1.044)	0.920
Free Cholesterol in Very Large HDL	1.006	(0.97 ,1.044)	0.866
Triglycerides in Very Large HDL	0.998	(0.968 ,1.03)	0.941
Concentration of Large HDL Particles	1.008	(0.969 ,1.048)	0.857
Total Lipids in Large HDL	1.012	(0.973 ,1.052)	0.732
Phospholipids in Large HDL	1.016	(0.977 ,1.057)	0.579
Cholesterol in Large HDL	1.007	(0.968 ,1.047)	0.872
Cholesteryl Esters in Large HDL	1.005	(0.966 ,1.046)	0.904
Free Cholesterol in Large HDL	1.011	(0.972 ,1.051)	0.750
Triglycerides in Large HDL	1.010	(0.978 ,1.043)	0.707
Concentration of Medium HDL Particles	1.001	(0.965 ,1.037)	0.981
Total Lipids in Medium HDL	1.003	(0.968 ,1.039)	0.934
Phospholipids in Medium HDL	1.005	(0.97 ,1.041)	0.902
Cholesterol in Medium HDL	1.001	(0.965 ,1.038)	0.981
Cholesteryl Esters in Medium HDL	1.000	(0.965 ,1.037)	0.983
Free Cholesterol in Medium HDL	1.002	(0.966 ,1.04)	0.941
Triglycerides in Medium HDL	1.003	(0.972 ,1.036)	0.925

Concentration of Small HDL Particles	0.972	(0.942 ,1.004)	0.216
Total Lipids in Small HDL	0.983	(0.951 ,1.015)	0.475
Phospholipids in Small HDL	0.988	(0.957 ,1.021)	0.639
Cholesterol in Small HDL	0.974	(0.944 ,1.006)	0.247
Cholesteryl Esters in Small HDL	0.972	(0.942 ,1.004)	0.214
Free Cholesterol in Small HDL	0.985	(0.953 ,1.017)	0.520
Triglycerides in Small HDL	0.998	(0.965 ,1.032)	0.936
Total Fatty Acids	0.982	(0.952 ,1.014)	0.439
Polyunsaturated Fatty Acids	0.960	(0.93 ,0.991)	0.057
Monounsaturated Fatty Acids	0.996	(0.965 ,1.029)	0.920
Saturated Fatty Acids	0.992	(0.961 ,1.024)	0.759
Linoleic Acid	0.950	(0.921 ,0.981)	0.057
Docosahexaenoic Acid	0.994	(0.962 ,1.027)	0.857
Omega-3 Fatty Acids	0.991	(0.96 ,1.024)	0.751
Omega-6 Fatty Acids	0.957	(0.927 ,0.988)	0.057
Degree of Unsaturation	0.976	(0.944 ,1.009)	0.284
Total Cholines	0.977	(0.944 ,1.011)	0.329
Phosphatidylcholines	0.982	(0.949 ,1.017)	0.482
Sphingomyelins	0.983	(0.95 ,1.016)	0.482
Phosphoglycerides	0.983	(0.951 ,1.017)	0.500
Alanine	0.976	(0.946 ,1.007)	0.265
Glutamine	0.997	(0.966 ,1.029)	0.925
Glycine	0.981	(0.946 ,1.016)	0.463
Histidine	0.976	(0.946 ,1.007)	0.265
Total Concentration of Branched-Chain Amino Acids (Leucine + Isoleucine + Valine)	1.009	(0.977 ,1.042)	0.751

Isoleucine	1.010	(0.979 ,1.042)	0.689
Leucine	1.006	(0.974 ,1.039)	0.866
Valine	1.009	(0.977 ,1.042)	0.750
Phenylalanine	1.026	(0.995 ,1.058)	0.227
Tyrosine	1.026	(0.995 ,1.058)	0.227
Glucose	1.022	(0.994 ,1.051)	0.248
Lactate	0.984	(0.954 ,1.016)	0.493
Pyruvate	0.995	(0.964 ,1.028)	0.902
Citrate	1.015	(0.983 ,1.048)	0.531
Acetone	1.047	(1.02 ,1.075)	0.056
Acetoacetate	1.043	(1.013 ,1.074)	0.057
Acetate	1.011	(0.986 ,1.036)	0.560
3-Hydroxybutyrate	1.048	(1.018 ,1.078)	0.056
Glycoprotein Acetyls	1.037	(1.003 ,1.072)	0.096
Creatinine	1.034	(0.998 ,1.071)	0.178
Albumin	0.928	(0.899 ,0.958)	<0.001*

CVD: cardiovascular disease; HR: hazard ratio; CI: confidential interval; FDR, false discovery rate; VLDL: Very-low-density lipoprotein; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Metabolites concentrations were standardized by z-score.

* indicates FDR-adjusted p values < 0.05.

Model adjusted for age at recruitment, sex, ethnicity, assessment centre, body mass index, Townsend deprivation index, household income, education, smoking status, alcohol drinking frequency, sleep duration, physical activity, and healthy diet score.