

RESEARCH ARTICLE

Obesity, metabolic factors and risk of different histological types of lung cancer: A Mendelian randomization study

Robert Carreras-Torres¹, Mattias Johansson¹, Philip C. Haycock², Kaitlin H. Wade², Caroline L. Relton², Richard M. Martin^{2,3}, George Davey Smith², Demetrius Albanes⁴, Melinda C. Aldrich⁵, Angeline Andrew⁶, Susanne M. Arnold⁷, Heike Bickeböllner⁸, Stig E. Bojesen⁹, Hans Brunnström¹⁰, Jonas Manjer¹⁰, Irene Brüske¹¹, Neil E. Caporaso¹², Chu Chen¹³, David C. Christiani¹⁴, W. Jay Christian¹⁵, Jennifer A. Doherty¹⁶, Eric J. Duell¹⁷, John K. Field¹⁸, Michael P. A. Davies¹⁸, Michael W. Marcus¹⁸, Gary E. Goodman¹⁹, Kjell Grankvist²⁰, Aage Haugen²¹, Yun-Chul Hong²², Lambertus A. Kiemeny²³, Erik H. F. M. van der Heijden²³, Peter Kraft²⁴, Mikael B. Johansson²⁰, Stephen Lam²⁵, Maria Teresa Landi¹², Philip Lazarus²⁶, Loïc Le Marchand²⁷, Geoffrey Liu²⁸, Olle Melander^{29,30}, Sungshim L. Park³¹, Gad Rennert³², Angela Risch³³, Eric B. Haura³⁴, Ghislaine Scelo¹, David Zaridze³⁵, Anush Mukeriyar³⁵, Milan Savic³⁶, Jolanta Lissowska³⁷, Beata Swiatkowska³⁸, Vladimir Janout³⁹, Ivana Holcatova⁴⁰, Dana Mates⁴¹, Matthew B. Schabath⁴², Hongbing Shen⁴³, Adonina Tardon⁴⁴, M Dawn Teare⁴⁵, Penella Woll⁴⁵, Ming-Sound Tsao⁴⁶, Xifeng Wu⁴⁷, Jian-Min Yuan⁴⁸, Rayjean J. Hung⁴⁹, Christopher I. Amos⁵⁰, James McKay¹, Paul Brennan^{1*}



OPEN ACCESS

Citation: Carreras-Torres R, Johansson M, Haycock PC, Wade KH, Relton CL, Martin RM, et al. (2017) Obesity, metabolic factors and risk of different histological types of lung cancer: A Mendelian randomization study. *PLoS ONE* 12(6): e0177875. <https://doi.org/10.1371/journal.pone.0177875>

Editor: Cheng Hu, Shanghai Diabetes Institute, CHINA

Received: April 20, 2017

Accepted: May 4, 2017

Published: June 8, 2017

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: RCT, MJ, PCH, KHW, CR, RMM, GDS, and PB are investigators or researchers on a Cancer Research UK (C18281/A19169) Programme Grant (the Integrative Cancer Epidemiology Programme). RMM is supported by the National Institute for Health Research (NIHR) Bristol Nutritional Biomedical Research Unit based

1 Section of Genetics, International Agency for Research on Cancer (IARC), Lyon, France, **2** MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom, **3** National Institute for Health Research Biomedical Research Unit in Nutrition, Diet and Lifestyle at University Hospitals Bristol NHS Foundation Trust and the University of Bristol, Bristol, United Kingdom, **4** Metabolic Epidemiology Branch, DCEG, National Cancer Institute, NIH, Rockville, Maryland, United States of America, **5** Department of Thoracic Surgery, Division of Epidemiology, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America, **6** Norris Cotton Cancer Center, Lebanon, New Hampshire, United States of America, **7** University of Kentucky Markey Cancer Center, Lexington, Kentucky, United States of America, **8** University Medical Center Göttingen, Göttingen, Germany, **9** Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, **10** Laboratory Medicine Region Skåne, Department of Clinical Sciences Lund, Lund University, Lund, Sweden, **11** Helmholtz Zentrum München, Munich, Germany, **12** Genetic Epidemiology Branch, DCEG, National Cancer Institute, NIH, Rockville, Maryland, United States of America, **13** Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, **14** Departments of Environmental Health and Epidemiology, Harvard T. H. Chan School of Public Health and Department of Medicine, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, United States of America, **15** Department of Epidemiology, College of Public Health, University of Kentucky, Lexington, Kentucky, United States of America, **16** Department of Epidemiology, Geisel School of Medicine, Dartmouth College, Lebanon, New Hampshire, United States of America, **17** Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology (ICO-IDIBELL), Barcelona, Spain, **18** Roy Castle Lung Cancer Research Programme, Department of Molecular and Clinical Cancer Medicine, Institute of Translational Medicine, The University of Liverpool, Liverpool, United Kingdom, **19** Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, **20** Department of Medical Biosciences, Umeå University, Umeå, Sweden, **21** National Institute of Occupational Health, Oslo, Norway, **22** Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea, **23** Radboud University Medical Center, Nijmegen, The Netherlands, **24** Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States of America, **25** British Columbia Cancer Agency, Vancouver, British Columbia, Canada, **26** Washington State University College of Pharmacy, Spokane, Washington, United States of America, **27** Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii, United States of America, **28** Ontario Cancer Institute, Princess Margaret Cancer Center, Toronto, Ontario, Canada, **29** Department of Clinical Sciences Malmö, Lund University, Lund, Sweden, **30** Department of Internal Medicine, Skåne University Hospital, Malmö, Sweden, **31** Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America, **32** Clalit National Cancer Control Center and Department of Community Medicine and Epidemiology, Carmel Medical Center and B. Rappaport Faculty of Medicine, Technion-Israel

at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. CLR and GDS are supported by funding from the MRC Integrative Epidemiology Unit at the University of Bristol (MC_UU_12013/1, MC_UU_12013/2). PCH is supported by a Cancer Research UK Population Research Postdoctoral Fellowship (C52724/A20138). SMA was supported by the Department of Defense under award number: 10153006 (W81XWH-11-1-0781) and by the UK Center for Clinical and Translational Science, (UL1TR000117). JMY is partially supported by the U.S. National Institutes of Health Grants (R01 CA144034 and UM1 CA182876). CARET investigators would like to thank the study participants for their involvement and acknowledge the National Cancer Institute and National Institute of Health for their grant support: 5-UM1-CA-167462, (PI Gary E. Goodman), U01-CA63673 (PIs G. Omenn, G. Goodman), R01-CA111703 (PI Chu Chen), and 5R01-CA151989-01A1 (PI Jennifer Doherty).

Competing interests: The authors have declared that no competing interests exist.

Institute of Technology, Haifa, Israel, **33** University of Salzburg and Cancer Cluster Salzburg, Salzburg, Austria, **34** Department of Thoracic Oncology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida, United States of America, **35** Russian N.N. Blokhin Cancer Research Centre, Moscow, The Russian Federation, **36** Department of Thoracic Surgery, Clinical Center of Serbia, Belgrade, Serbia, **37** Department of Cancer Epidemiology and Prevention, Maria Skłodowska-Curie Institute – Oncology Center, Warsaw, Poland, **38** Nofer Institute of Occupational Medicine, Department of Environmental Epidemiology, Lodz, Poland, **39** Department of Epidemiology and Public Health, Faculty of Medicine, University of Ostrava, Ostrava, Czech Republic, **40** Institute of Public Health and Preventive Medicine, Charles University, 2nd Faculty of Medicine, Prague, Czech Republic, **41** National Institute of Public Health, Bucharest, Romania, **42** Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida, United States of America, **43** Department of Epidemiology and Biostatistics, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, Nanjing, China, **44** Instituto Universitario de Oncología del Principado de Asturias (IUOPA), University of Oviedo and CIBERESP, Oviedo, Spain, **45** University of Sheffield, Sheffield, United Kingdom, **46** Princess Margaret Cancer Center, Toronto, Canada, **47** The University of Texas MD Anderson Cancer Center, Houston, Texas, United States of America, **48** Division of Cancer Control and Population Science, University of Pittsburgh Cancer Institute; and Department of Epidemiology, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania, United States of America, **49** Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada, **50** Department of Biomedical Data Science, Geisel School of medicine, Dartmouth College, Lebanon, New Hampshire, United States of America

* gep@iarc.fr

Abstract

Background

Assessing the relationship between lung cancer and metabolic conditions is challenging because of the confounding effect of tobacco. Mendelian randomization (MR), or the use of genetic instrumental variables to assess causality, may help to identify the metabolic drivers of lung cancer.

Methods and findings

We identified genetic instruments for potential metabolic risk factors and evaluated these in relation to risk using 29,266 lung cancer cases (including 11,273 adenocarcinomas, 7,426 squamous cell and 2,664 small cell cases) and 56,450 controls. The MR risk analysis suggested a causal effect of body mass index (BMI) on lung cancer risk for two of the three major histological subtypes, with evidence of a risk increase for squamous cell carcinoma (odds ratio (OR) [95% confidence interval (CI)] = 1.20 [1.01–1.43] and for small cell lung cancer (OR [95%CI] = 1.52 [1.15–2.00]) for each standard deviation (SD) increase in BMI [4.6 kg/m²]), but not for adenocarcinoma (OR [95%CI] = 0.93 [0.79–1.08]) ($P_{\text{heterogeneity}} = 4.3 \times 10^{-3}$). Additional analysis using a genetic instrument for BMI showed that each SD increase in BMI increased cigarette consumption by 1.27 cigarettes per day ($P = 2.1 \times 10^{-3}$), providing novel evidence that a genetic susceptibility to obesity influences smoking patterns. There was also evidence that low-density lipoprotein cholesterol was inversely associated with lung cancer overall risk (OR [95%CI] = 0.90 [0.84–0.97] per SD of 38 mg/dl), while fasting insulin was positively associated (OR [95%CI] = 1.63 [1.25–2.13] per SD of 44.4 pmol/l). Sensitivity analyses including a weighted-median approach and MR-Egger test did not detect other pleiotropic effects biasing the main results.

Conclusions

Our results are consistent with a causal role of fasting insulin and low-density lipoprotein cholesterol in lung cancer etiology, as well as for BMI in squamous cell and small cell carcinoma. The latter relation may be mediated by a previously unrecognized effect of obesity on smoking behavior.

Introduction

Lung cancer is the leading cause of cancer mortality [1]. Most lung cancers are caused by tobacco smoking [2], although associations have also been reported with a range of inflammatory and metabolic conditions. Observational studies have indicated an inverse relationship for both body mass index (BMI) [3–7] and lipid levels [8,9], as well as a positive correlation with dietary glycemic index [10] and insulin levels [11], with lung cancer risk. However, given the strong effect of tobacco smoking on lung cancer risk, and the well described association between tobacco consumption and alterations in body weight [12–14], traditional observational studies are unlikely to fully account for the confounding effect of tobacco exposure when describing the relationship between lung cancer and obesity or metabolic conditions.

Mendelian randomization (MR) is an analytical approach based on genetic markers of an exposure (i.e. an instrumental variable) that is less sensitive to reverse causation and confounding than traditional regression analyses in observational studies [15]. Inherited gene variants associated with the risk factors of interest should act as unconfounded markers of those risk factors, assuming an absence of pleiotropy [16]. In this instance, an association between the genetic variant and the outcome implies that the risk factor of interest may have a causal effect on the outcome [17]. Previous MR analyses on lung cancer risk showed that a genetic score for increased BMI raised the risk for lung cancer, especially for squamous cell carcinoma and small cell lung cancer [18,19]. These results are in contradiction to observational findings, confirming the utility of MR analyses for this exposure.

The goal of the current study was to use genetic variations associated with a range of metabolic factors, including obesity, body shape, dyslipidemia and hyperglycemia, to further investigate the causal relationship between these metabolic exposures and lung cancer. The method applied in this study is called two-sample MR, which combines summary statistics on genetic variant-exposure and genetic variant-outcome associations from different samples [20,21]. Furthermore, we sought to confirm that the genetic control of these metabolic phenotypes did not influence cigarette smoking behavior.

Methods

Genetic instruments for obesity and metabolic parameters

Genetic instruments for potential risk factors were independent (linkage disequilibrium (LD) $R^2 < 0.01$ in European 1000 Genomes Phase3 samples [22]) single nucleotide polymorphisms (SNPs) that were associated with the trait of interest in the most recent genome-wide association study (GWAS) ($P < 5 \times 10^{-8}$) on European ancestry samples. Results from the Genetic Investigation of ANthropometric Traits (GIANT) consortium were used to identify genetic proxies for body mass index (BMI) [23], and waist to hip ratio [24]. High-density and low-density lipoprotein cholesterol (HDL and LDL), total cholesterol and triglycerides were selected as lipid profile components. Genetic loci influencing bloodstream levels of these lipids were

Table 1. Number of identified instrumental SNPs for metabolic factors, phenotype distribution in the discovery sample, and proportion of phenotype variance explained by the instruments. SD: standard deviation.

Phenotype	N SNP	Mean ± SD	Units	Variance explained (%)	Consortium	Publication for instruments
Body mass index	73	27.0 ± 4.6	kg/m ²	2.7	GIANT	Locke et al 2015 [23]
Waist to hip ratio	32	1.1 ± 0.1	cm/cm	1.4	GIANT	Shungin et al 2015 [24]
High density cholesterol	63	53.3 ± 15.5	mg/dl	13.7	GLGC	Willer et al 2013 [25]
Low density cholesterol	50	133.6 ± 38.0	mg/dl	14.6	GLGC	Willer et al 2013 [25]
Total cholesterol	67	213.28 ± 42.6	mg/dl	15.0	GLGC	Willer et al 2013 [25]
Triglycerides	39	140.85 ± 87.8	mg/dl	11.7	GLGC	Willer et al 2013 [25]
HDL rare variants	11	1.4 ± 0.4	mmol/l	4.1	-	Helgadottir et al 2016 [26]
Non-HDL rare variants	11	4.0 ± 1.2	mmol/l	4.1	-	Helgadottir et al 2016 [26]
Triglycerides rare variants	7	1.3 ± (0.8–2.2)	mmol/l	2.4	-	Helgadottir et al 2016 [26]
Fasting glucose	25	5.2 ± 0.8	mmol/l	4.8	MAGIC	Scott et al 2012 [27]
Fasting insulin	11	56.9 ± 44.4	pmol/l	1.2	MAGIC	Scott et al 2012 [27]
2h post-challenge glucose	6	5.6 ± 1.7	mmol/l	1.7	MAGIC	Scott et al 2012 [27]

<https://doi.org/10.1371/journal.pone.0177875.t001>

identified from GWA data provided by the Global Lipids Genetic Consortium (GLGC) [25]. Additionally, genetic instruments of low frequency were also identified for HDL, LDL cholesterol and triglycerides [26]. Finally, hyperglycemia and hyperinsulinemia parameters were reported within the Meta-Analysis of Glucose and Insulin related traits Consortium (MAGIC) at fasting levels and during an oral glucose tolerance test (OGTT). Genetic instruments were identified for plasma levels of fasting glucose, fasting insulin and 2-hour post-challenge glucose [27]. For these metabolic parameters, in order to reduce potential pleiotropic effect from BMI, SNPs that were not associated with the phenotype once adjusted for BMI ($P > 0.05$) were excluded from our analyses. For each identified SNP, the reported effect was for the increasing trait allele expressed in standard deviations (SD) of the trait per-allele (β_{GP}) along with the standard error (SE_{GP}). For the studies in which the genetic effects were not originally reported in SDs of the trait, they were recalibrated according to the mean SD, weighted for sample size, reported across different case-control samples. SNPs with ambiguous strand codification (A/T or C/G) were replaced by SNPs in tight genetic linkage ($R^2 > 0.8$) using the *proxysnps* R package (European populations) (R Project) or removed from the analyses. Details on i) the number of SNPs identified for each potential risk trait, ii) the mean and SD of the traits in the discovery GWAS, and iii) the proportion of trait variance explained by the genetic proxies are presented in Table 1.

Lung cancer association results

GWAS results on overall lung cancer risk and by histology type were obtained from the recently completed meta-analyses [28] comprising the previous TRICL-ILCCO lung cancer GWAS (11,177 lung cancer cases and 40,396 controls) [29] and an additional 18,089 lung cancer samples and 16,054 controls genotyped using the Illumina Infinium OncoArray-500K BeadChip (Illumina Inc. San Diego, CA). The overall sample size was 29,266 lung cancer cases and 56,450 controls. GWAS analyses were stratified by histology, including 11,273 adenocarcinomas, 7,426 squamous cell carcinomas and 2,664 small cell lung cancers. Additionally, analyses were performed stratified by smoking status defined as ever smokers (current and former smokers; 23,223 cases and 16,964 controls) and never smokers (2,355 cases and 7,504 controls). SNPs with imputation quality score (R^2) less than 0.7 were removed from the datasets. For all SNPs used as instruments in this report, SNP to phenotype effect (β_{GP}) and SNP to disease effect (β_{GD}) for each lung cancer subgroup can be observed in S1 Table.

Power assessment

Power calculations on MR analyses were performed according to Burgess [30], given the number of cases and controls of the lung cancer subgroups, for five genetic instruments accounting for different proportion of explained variance (1%, 2.5%, 5%, 10% and 15%) within the range of our genetic instrument sets (see Table 1). Power calculations are presented in S1 Fig.

Mendelian randomization analyses on lung cancer

The causal risk effects of metabolic factors on lung cancer overall, as well as by histology and other subgroups were estimated using a likelihood-based approach, which is considered the most accurate method to estimate causal effects when there is a continuous log-linear association between risk factor and disease risk [20]. The resulting odds ratio (OR) and 95% confidence interval (CI) provide an estimate of relative risk caused by each SD increase in the exposure. We investigated the heterogeneity of causal effects by histology and by smoking status, estimating the I^2 statistic assuming a fixed-effect model using the *meta* R package (R project). We also provide OR estimates using the complementary weighted median method wherein the causal effect estimate is weighted towards the median of the distribution of instrumental SNPs [31]. This approach is less sensitive to individual instrumental SNPs behaving as outliers. Furthermore, because OR estimates can be confounded if the instrumental genetic variants are also associated with other causal risk factors, i.e. through directional pleiotropy, we also performed MR-Egger [32]. This approach involves a weighted linear regression of the SNP to disease effects (β_{GD}) on the SNP to phenotype effects (β_{GP}), and provides a direct estimate of any directional pleiotropy under the assumption that pleiotropic biases affecting the overall MR risk estimates are acting in the same direction [33]. Additionally, MR-Egger regression is able to identify SNPs behaving as genetic outliers, providing Bonferroni corrected P-values. We also provide a funnel plot of the fold risk increase ($\exp(\beta_{GD}/\beta_{GP})$) vs. the instrumental strength of each SNP (β_{GP}/SE_{GD}^2) to allow visual assessment of asymmetry of instrumental causal estimates. These plots were generated using the *ggplot2* R package (R Project).

Assessment of association between the genetic instruments and other phenotypes

To evaluate whether MR lung cancer risk estimates could be generated by other phenotypes than the instrumented ones, the association of genetic instruments with other phenotypes was assessed. The effects of our genetic instruments on cigarette smoking was evaluated using a novel online platform for Mendelian randomization based on summary genetic association data from GWAS (MR-Base; <http://www.mrbase.org/>) [34]. Genetic effects on cigarette smoking (cigarettes per day) were obtained from the Tobacco and Genetics (TAG) consortium [35]. TAG data were generated from 16 independent studies which included 38,181 smokers with information on number of smoked cigarettes per day. We also assessed the association of dyslipidemia and hyperglycemia genetic instruments with BMI using data from the GIANT consortium [23].

Results

Obesity and body shape parameters

Based on 72 SNPs used as genetic instruments for BMI, we estimated an overall OR (per SD increase in BMI (4.6 kg/m^2)) of 1.07 (95%CI = 0.96–1.20) for lung cancer risk (Fig 1). However, stratified analyses by histology subtypes revealed heterogeneity in the causal effect estimates ($P_{\text{heterogeneity}} = 4.3 \times 10^{-3}$), with no effect on adenocarcinoma (OR [95%CI] = 0.93 [0.79–

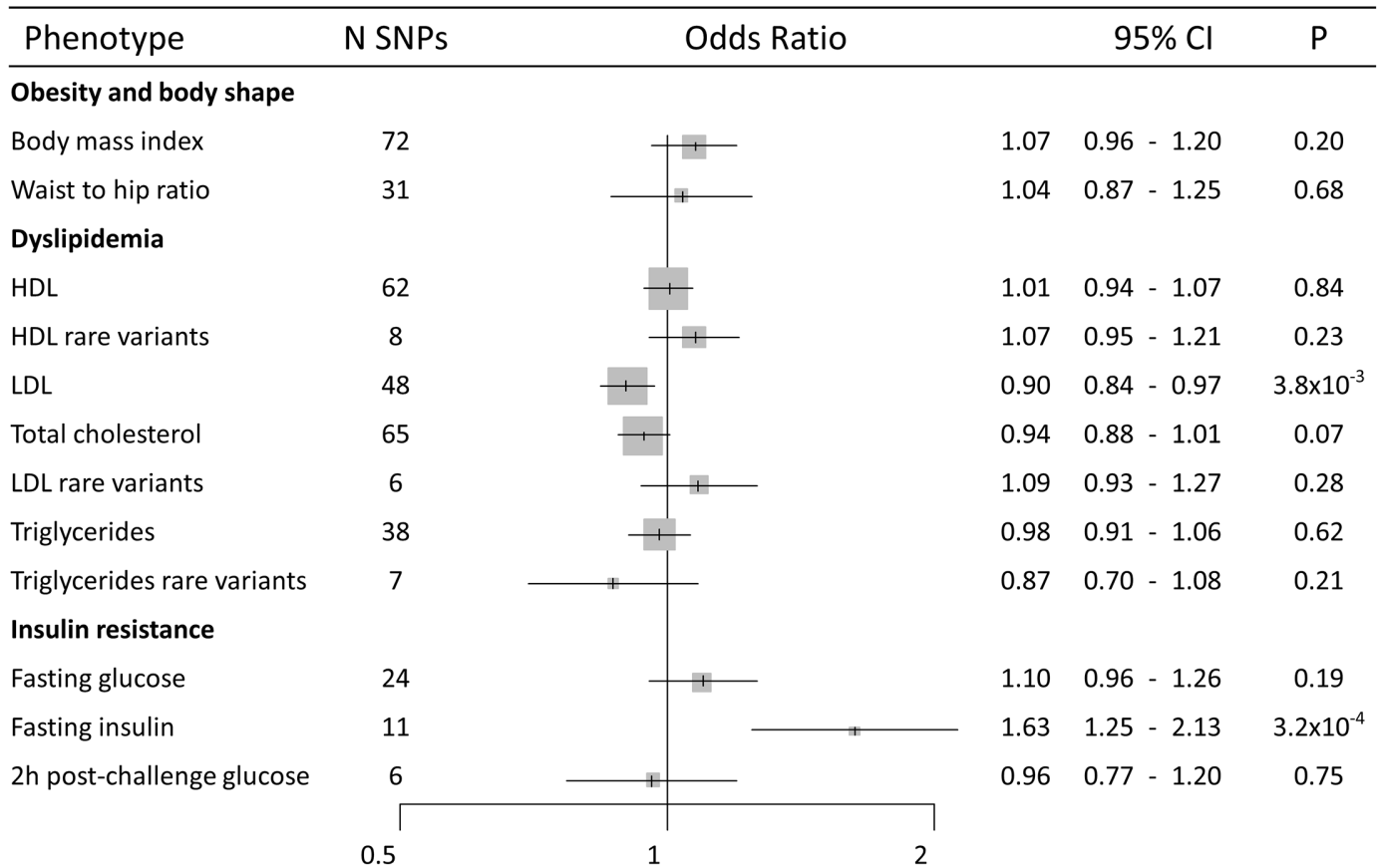


Fig 1. Forest plot of lung cancer overall risk for one standard deviation (Table 1) increase in each phenotype provided by the MR likelihood-based approach. CI: Confidence interval. P: P value.

<https://doi.org/10.1371/journal.pone.0177875.g001>

1.08]), but an increased risk for squamous cell carcinoma (OR [95%CI] = 1.20 [1.01–1.43]) and small cell lung cancer (OR [95%CI] = 1.52 [1.15–2.00]) (Fig 2). Analyses stratified by smoking status showed a null effect of BMI in never smokers (OR [95%CI] = 0.83 [0.60–1.16]) and in ever smokers (OR [95%CI] = 1.12 [0.97–1.29]) ($P_{heterogeneity} = 0.10$) (Fig 2). Weighted median MR approach provided similar risk estimates for the different lung cancer types (S2 Table). The MR-Egger intercept test revealed the presence of bias on the initial risk estimation on lung cancer overall and in ever-smokers (S3 Table), with MR-Egger risk estimates indicating no effect of BMI on lung cancer overall. Potential asymmetry of the SNP risk estimates reflected the presence of directional pleiotropy for lung cancer overall (Fig 3A). Funnel plots for the other histological types can be observed in the funnel plots included in Fig 3B–3F. Regarding waist-to-hip ratio, there was no evidence of causal effect on lung cancer overall (OR [95%CI] = 1.04 [0.87–1.25]) (Fig 1) or histology subgroups (S2 Fig). Sensitivity analyses did not detect any effect biasing the initial risk estimates (S2 and S3 Tables). Funnel plots can be observed in S3 Fig.

Blood lipid levels

There was no consistent evidence of a causal effect of HDL on lung cancer overall (OR [95%CI] = 1.01 [0.94–1.07] for instrumental common SNPs, and OR [95%CI] = 1.07 [0.95–1.21] for instrumental rare SNPs) (Fig 1), nor any lung cancer subgroup (S4–S7 Figs).

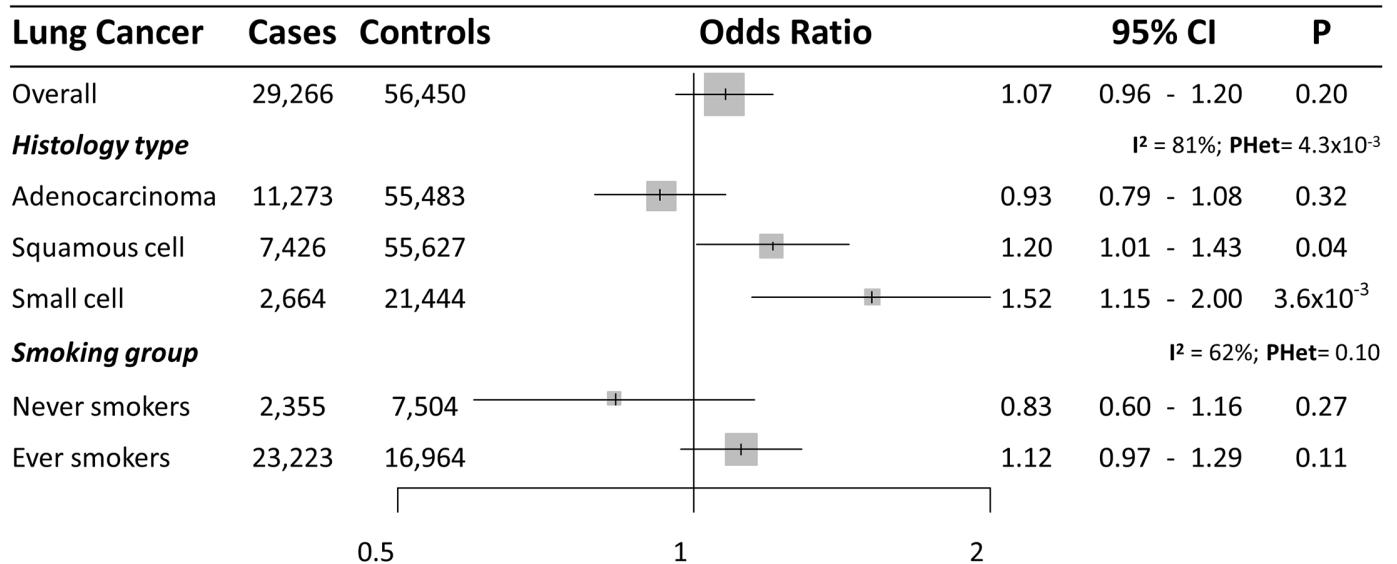


Fig 2. Forest plot of lung cancer risk for each SD increase in BMI (approximately 4.6 kg/m²) observed in the likelihood-based MR approach. 95%CI: 95% Confidence Interval; P: P value. *I*²: between-strata heterogeneity. *P*Het: P value of between-strata heterogeneity.

<https://doi.org/10.1371/journal.pone.0177875.g002>

Triglyceride levels did not indicate any effect on lung cancer overall; (OR [95%CI] = 0.98 [0.91–1.06] for instrumental common SNPs and OR [95%CI] = 0.87 [0.70–1.08] for instrumental rare SNPs) (Fig 1), or for lung cancer subtypes (S8–S11 Figs). An inverse relationship was observed with overall lung cancer risk for each SD increase in LDL (approximately 38.0 mg/dl) (OR [95%CI] = 0.90 [0.84–0.97] for instrumental common SNPs (Fig 1), although stratified analyses did not provided a consistent inverse association among histology types (OR of 0.93 for adenocarcinoma (95%CI = 0.85–1.03), 0.88 for squamous cell carcinoma (95%CI = 0.79–0.98), and 0.96 for small cell lung cancer (95%CI = 0.81–1.14) (Fig 4 and S12 Fig for funnel plots). Additionally, LDL instrumental rare SNPs did not provide evidence of association with lung cancer overall (OR [95%CI] = 1.09 [0.93–1.27]) (Fig 1) or subtypes (S13 and S14 Figs for funnel plots). Finally, total cholesterol showed a similar pattern of association with lung cancer overall and with subtypes that LDL common SNPs (OR of 0.94 for lung cancer overall (95% CI = 0.88–1.01), OR of 1.01 for adenocarcinoma (95%CI = 0.92–1.10), 0.89 for squamous cell carcinoma (95%CI = 0.80–1.00), and 0.93 for small cell lung cancer (95%CI = 0.79–1.10)) (S15 and S16 Figs for funnel plots).

Insulin resistance parameters

Among glucose and insulin parameters, fasting insulin was associated with an increased risk in overall lung cancer (OR [95%CI] = 1.63 [1.25–2.13] per each SD increase [44.4 pmol/l]) (Fig 1). Stratified analyses provided consistent associations with all histology subtypes (Fig 5), with a risk increase of 1.60 for adenocarcinoma (95%CI = 1.12–2.31), 1.84 for squamous cell carcinoma (95%CI = 1.20–2.81), and 2.46 for small cell lung cancer (95%CI = 1.26–4.81). Stratified analyses by smoking status showed some heterogeneity in causal effects (*P*_{heterogeneity} = 0.16) (Fig 5), with a clear increased risk among ever smokers (OR [95%CI] = 1.86 [1.33–2.60]), but no association among never smokers (OR [95%CI] = 1.01 [0.46–2.20]). Sensitivity tests identified rs4865796 as an outlier SNP in the analyses for lung adenocarcinoma subtype and lung cancer overall in ever smokers. Removing this SNP, the analyses indicated slightly higher levels of risk associated with this genetic instrument (Funnel plots in S17 Fig).

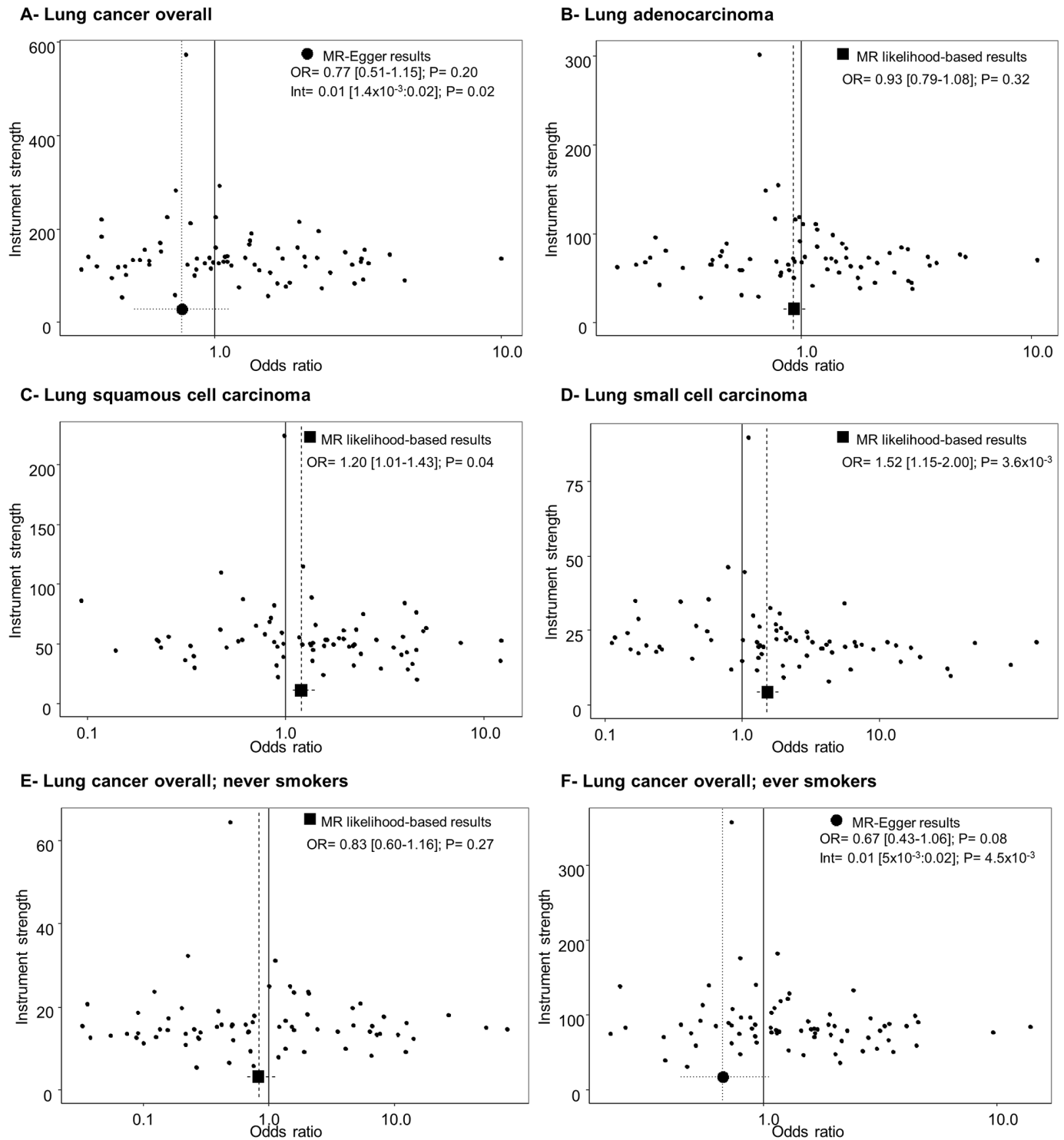


Fig 3. Funnel plots for the distribution of risk estimates of BMI instrumental SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

<https://doi.org/10.1371/journal.pone.0177875.g003>

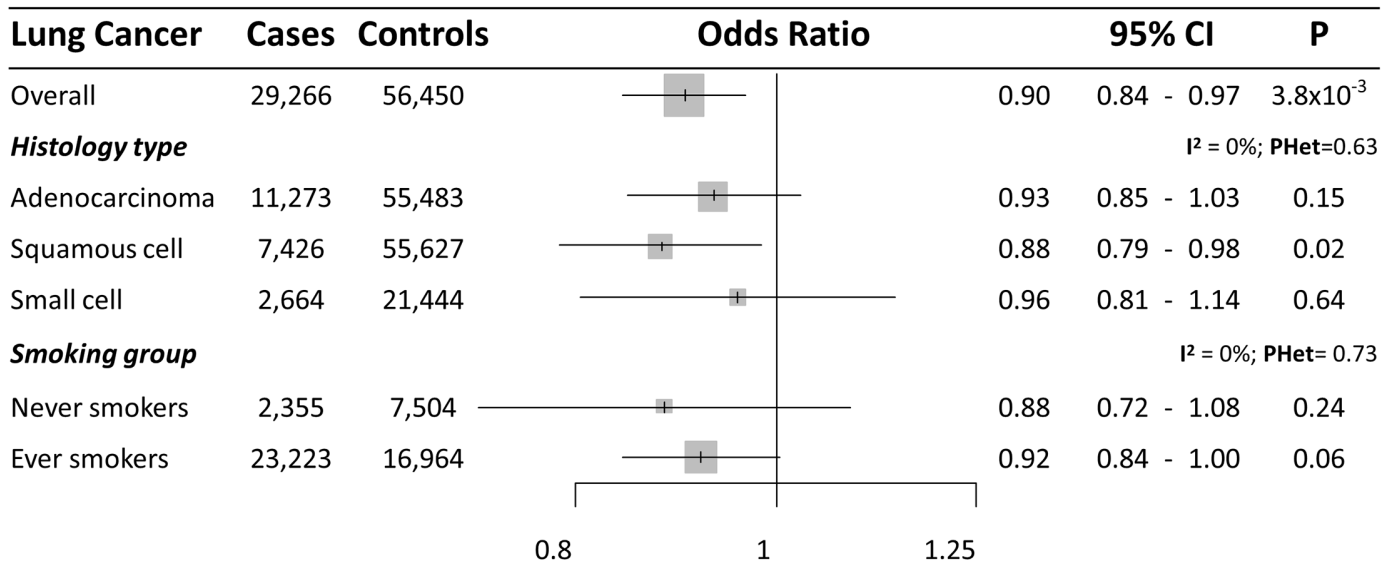


Fig 4. Forest plot of lung cancer risk for each SD increase in LDL (approximately 38.0 mg/dl) observed in the likelihood-based MR approach using the instrument set of common SNPs. 95%CI: 95% Confidence Interval; P: P value. *I*²: between-strata heterogeneity. *P*Het: P value of between-strata heterogeneity.

<https://doi.org/10.1371/journal.pone.0177875.g004>

Fasting glucose showed little evidence for an association with overall lung cancer risk (OR per SD of 0.8 mmol/l = 1.10; 95%CI = 0.96–1.26) (Fig 1), nor for adenocarcinoma (OR [95% CI] = 1.03 [0.85–1.24] or small cell lung cancer (OR [95%CI] = 1.19 [0.83–1.68]), where a positive association with risk was seen for squamous cell carcinoma (OR [95%CI] = 1.32 [1.06–1.65]) (S18 and S19 Figs for funnel plots). Finally, 2-hour post-challenge glucose showed no evidence of causality for lung cancer overall (OR [95%CI] = 0.96 [0.77–1.20]) (Fig 1) or any lung cancer subgroup (S20 and S21 Figs for funnel plots).

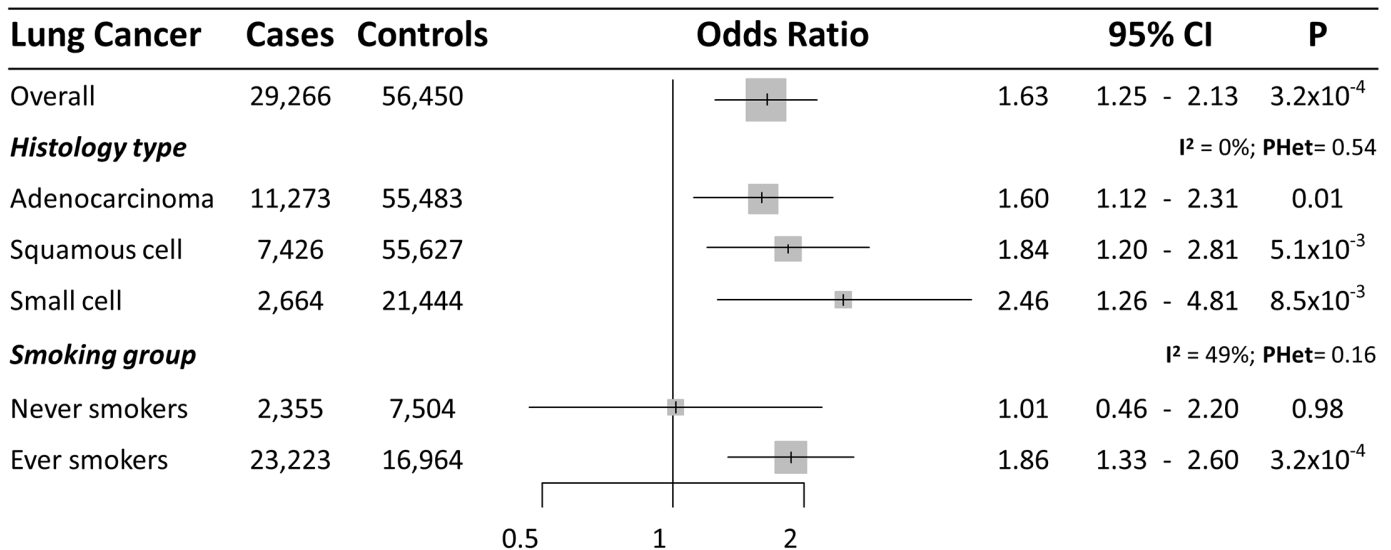


Fig 5. Forest plot of lung cancer risk for each SD increase in fasting insulin (approximately 44.4 pmol/L) observed in the likelihood-based MR approach. 95%CI: 95% Confidence Interval; P: P value. *I*²: between-strata heterogeneity. *P*Het: P value of between-strata heterogeneity.

<https://doi.org/10.1371/journal.pone.0177875.g005>

Assessment of association between the genetic instruments and other phenotypes

Using MR-Base, we also evaluated the association between cigarette smoking for each genetic instrument associated with lung cancer risk, including BMI, LDL, and fasting insulin. Each SD increase in BMI increased cigarette consumption by 1.27 cigarettes (95%CI = 0.46:2.07; $P = 2.1 \times 10^{-3}$; MR likelihood-based approach), while the other metabolic genetic instruments were not associated with cigarette smoking (MR likelihood-based estimate per SD increase [95%CI] = -0.35 [-0.92:0.22] for LDL and 1.21 [-0.67:3.08] for fasting insulin). Additionally, each SD increase in LDL was inversely associated with BMI (-0.02 SD in BMI (95%CI = -0.04:-3.8 $\times 10^{-3}$; $P = 0.02$), while each SD increase in fasting was not associated with BMI (0.07 SD in BMI (95%CI = -0.01:0.15; $P = 0.10$)).

Discussion

We have used existing data from large GWAS on metabolic factors to identify potentially causal risk factors of lung cancer that are amenable to intervention. Our results suggested that higher BMI is associated with an increase in risk of squamous and small cell carcinoma of the lung, and that this relation may be mediated by altered smoking behavior. Additionally, we found consistent evidence of a positive association between fasting insulin and overall lung cancer risk, as well as an inverse relationship between LDL levels and lung cancer risk.

Obesity and related metabolic factors are associated with risk of several cancers [36], but traditional observational studies are not reliable for assessing the association between metabolic factors and cancers strongly associated with smoking. The strength of the association between BMI and lung cancer subtypes found in this study, were weaker than those reported in previous MR analyses (45%-81% of risk increase) [18,19]. These previous studies were based on smaller sample sizes (12,160 and 16,572 lung cancer cases, respectively), whereas our current study provides causal estimates from 29,266 lung cancer cases and 56,450 controls, including the samples of previous studies. Our analysis also detected a strong positive effect of each unit increase in BMI on number of cigarettes smoked per day, indicating that the associations with risk of squamous and small cell carcinoma lung cancer are caused by an influence of BMI on smoking behavior. Given that this association was derived from genetic data on BMI, and therefore not influenced by any inverse effect that tobacco consumption may have on BMI levels, it has important implications beyond this study on the smoking behavior and smoking prevention.

Our results lend further support to epidemiological studies showing an increased lung cancer risk for the effect of reduced lipid levels [9] and elevated circulating insulin [11]. Several studies have reported an inverse relationship between total cholesterol levels and cancer risk [37], which has been previously interpreted as an effect of preclinical cancer due to an increased uptake of cholesterol by tumor cells [9]. In contrast, our study suggests a potential relationship between cholesterol levels and lung cancer. However, cholesterol genetic instruments using common and rare SNPs did not provide homogeneous results. This could be explained by different observations: by the inverse relationship between the instrument set of common SNPs for cholesterol and BMI revealed by a sensitivity analysis using MR-Base, by a multiple testing effect, but also by the different proportion of exposure that genetic instruments represent (instrument sets of rare SNPs were weaker and had less power to detect potential causal associations). Further MR mediation analyses are needed in order to clarify the causal relationship between lipid levels and lung cancer accounting for the potential effect of BMI and smoking. Finally, the risk increase observed by fasting glucose and fasting insulin is unlikely to be driven by cigarette smoking or BMI, since the relevant instrumental sets are not

associated with smoking or BMI. A potential causal relationship between increased insulin and cancer risk using genetic instruments has also been reported for endometrial cancer [38] and pancreatic cancer [39]. These results lend further support for an important role of the insulin-pathway in cancer aetiology [40–42].

One of the main caveats in Mendelian randomization studies is the potential violation of methodological assumptions, in particular the lack of directional pleiotropic effects biasing association estimates. In the current study we evaluated the extent to which pleiotropy may have affected our risk estimates using MR-Egger regression, as well as the MR-Base platform to directly assess the relation between genetic instruments and potential confounders.

In conclusion, using a two-sample Mendelian randomization approach, this study assessed a range of metabolic factors in relation to lung cancer risk. Our results revealed that genetic associations between BMI and lung cancer subtypes could be mediated by a positive relationship between BMI and cigarette consumption. Our results are also consistent with observational studies that indicate an inverse relationship between lipid levels and lung cancer risk, and an increasing risk caused by elevated fasting insulin levels.

Supporting information

S1 Fig. Power calculations for Mendelian randomization analyses on lung cancer using genetic instruments accounting for different proportion of phenotypic variance (15.0, 10.0, 5.0, 2.5, and 1.0%).

(PDF)

S2 Fig. Forest plot of lung cancer risk for each SD increase of waist-to-hip ratio observed in a likelihood-based MR approach. 95%CI: 95% Confidence Interval; P: P value. I^2 : between-strata heterogeneity. PHet: P value of between-strata heterogeneity.

(PDF)

S3 Fig. Funnel plots for the distribution of risk estimates of waist-to-hip ratio-instrumental SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S4 Fig. Forest plot of lung cancer risk for each SD increase in HDL observed in a likelihood-based MR approach using the instrument set of common SNPs. 95%CI: 95% Confidence Interval; P: P value. I^2 : between-strata heterogeneity. PHet: P value of between-strata heterogeneity.

(PDF)

S5 Fig. Forest plot of lung cancer risk for each SD increase in HDL observed in a likelihood-based MR approach using the instrument set of rare SNPs. 95%CI: 95% Confidence Interval; P: P value. I^2 : between-strata heterogeneity. PHet: P value of between-strata heterogeneity.

(PDF)

S6 Fig. Funnel plots for the distribution of risk estimates of HDL instrumental common SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S7 Fig. Funnel plots for the distribution of risk estimates of HDL instrumental rare SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S8 Fig. Forest plot of lung cancer risk for each SD increase in triglycerides observed in a likelihood-based MR approach using the main instrument set of common SNPs. 95%CI: 95% Confidence Interval; P: P value. I^2 : between-strata heterogeneity. PHet: P value of between-strata heterogeneity.

(PDF)

S9 Fig. Forest plot of lung cancer risk for each SD increase in triglycerides observed in a likelihood-based MR approach using the instrument set of rare SNPs. 95%CI: 95% Confidence Interval; P: P value. I^2 : between-strata heterogeneity. PHet: P value of between-strata heterogeneity.

(PDF)

S10 Fig. Funnel plots for the distribution of risk estimates of triglycerides instrumental common SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S11 Fig. Funnel plots for the distribution of risk estimates of triglycerides instrumental rare SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S12 Fig. Funnel plots for the distribution of risk estimates of LDL instrumental common SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S13 Fig. Forest plot of lung cancer risk for each SD increase in LDL observed in a likelihood-based MR approach using the instrument set of rare SNPs. 95%CI: 95% Confidence Interval; P: P value. I^2 : between-strata heterogeneity. PHet: P value of between-strata heterogeneity.

(PDF)

S14 Fig. Funnel plots for the distribution of risk estimates of LDL instrumental rare SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S15 Fig. Forest plot of lung cancer risk for each SD increase in total cholesterol observed in a likelihood-based MR approach. 95%CI: 95% Confidence Interval; P: P value. I^2 : between-strata heterogeneity. PHet: P value of between-strata heterogeneity.

(PDF)

S16 Fig. Funnel plots for the distribution of risk estimates of total cholesterol instrumental SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S17 Fig. Funnel plots for the distribution of risk estimates of fasting insulin instrumental SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S18 Fig. Forest plot of lung cancer risk for each SD increase in fasting glucose observed in a likelihood-based MR approach. 95%CI: 95% Confidence Interval; P: P value. I^2 : between-

strata heterogeneity. PHet: P value of between-strata heterogeneity.
(PDF)

S19 Fig. Funnel plots for the distribution of risk estimates of fasting glucose instrumental SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S20 Fig. Forest plot of lung cancer risk for each SD increase in glucose 2h post-challenge observed in a likelihood-based MR approach. 95%CI: 95% Confidence Interval; P: P value. I^2 : between-strata heterogeneity. PHet: P value of between-strata heterogeneity.

(PDF)

S21 Fig. Funnel plots for the distribution of risk estimates of glucose 2h post-challenge instrumental SNPs along with MR causal effect lung cancer subtypes. OR: Odds ratio; Int: Intercept; P: P value.

(PDF)

S1 Table. Association parameters of instrumental SNPs for the corresponding metabolic factor and for different lung cancer groups. CHR: Chromosome. BP: Base pair. SE: Standard error. BMI: Body mass index. HDL: High-density lipoprotein, LDL: Low-density lipoprotein.

(PDF)

S2 Table. Risk increase on lung cancer phenotypes for each standard deviation increase in the phenotype provided by weighted median MR approach. HDL: High-density lipoprotein, LDL: Low-density lipoprotein. Chol: Cholesterol. OR: Odds ratio. LCI: Lower confidence interval. UCI: Upper confidence interval. P: P value.

(PDF)

S3 Table. Overall pleiotropic effect assessment of causal estimates of potential risk factors on lung cancer phenotypes provided by MR-Egger test. Int: Intercept. HDL: High-density lipoprotein, LDL: Low-density lipoprotein. Chol: Cholesterol. OR: Odds ratio. Est: Estimate. LCI: Lower confidence interval. UCI: Upper confidence interval. P: P value.

(PDF)

Author Contributions

Conceptualization: MJ JMK PB.

Data curation: RCT.

Formal analysis: RCT.

Funding acquisition: MJ JMK PB PCH KHW CLR RMM GDS.

Investigation: DA MCA AA SMA H. Bickeböllner SEB H. Brunnström JM IB NEC CC DCC WJC JAD EJD JKF MPAD MWM GEG KG AH YCH LAK EHFMH PK MBJ SL MTL PL LLM GL OM SLP GR AR EBH MBS HS AT MDT PW MST XW JMY RJH CIA GS DZ AM MS JL BS VJ IH DM.

Methodology: PCH KHW CLR RMM GDS.

Project administration: MJ JMK PB.

Supervision: MJ JMK PB.

Writing – original draft: RCT MJ JMK PB.

Writing – review & editing: PCH KHW CLR RMM GDS DA MCA AA SMA H. Bickeböllner SEB H. Brunnström JM IB NEC CC DCC WJC JAD EJD JKF MPAD MWM GEG KG AH YCH LAK EHFMH PK MBJ SL MTL PL LLM GL OM SLP GR AR EBH MBS HS AT MDT PW MST XW JMY RJH CIA GS DZ AM MS JL BS VJ IH DM.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J. Global Cancer Statistics, 2012. *CA Cancer J Clin.* 2015; 65: 87–108. <https://doi.org/10.3322/caac.21262> PMID: 25651787
2. International Agency for Research on Cancer. List of Classifications by cancer sites with sufficient or limited evidence in humans, Volumes 1 to 114. [Internet]. 2015. <http://monographs.iarc.fr/ENG/Classification/index.php>
3. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon D a., Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. *Lancet.* 2014; 755–765. [https://doi.org/10.1016/S0140-6736\(14\)60892-8](https://doi.org/10.1016/S0140-6736(14)60892-8) PMID: 25129328
4. Kabat GC, Kim M, Hunt JR, Chlebowski RT, Rohan TE. Body mass index and waist circumference in relation to lung cancer risk in the women’s health initiative. *Am J Epidemiol.* 2008; 168: 158–169. <https://doi.org/10.1093/aje/kwn109> PMID: 18483121
5. Lam TK, Moore SC, Brinton L a., Smith L, Hollenbeck AR, Gierach GL, et al. Anthropometric Measures and Physical Activity and the Risk of Lung Cancer in Never-Smokers: A Prospective Cohort Study. *PLoS One.* 2013; 8. <https://doi.org/10.1371/journal.pone.0070672> PMID: 23940620
6. Smith L, Brinton L a., Spitz MR, Lam TK, Park Y, Hollenbeck AR, et al. Body mass index and risk of lung cancer among never, former, and current smokers. *J Natl Cancer Inst.* 2012; 104: 778–789. <https://doi.org/10.1093/jnci/djs179> PMID: 22457475
7. Yang Y, Dong J, Sun K, Zhao L, Zhao F, Wang L, et al. Obesity and incidence of lung cancer: A meta-analysis. *Int J Cancer.* 2013; 132: 1162–1169. <https://doi.org/10.1002/ijc.27719> PMID: 22777722
8. Kucharska-newton AM, Rosamond WD, Schroeder JC, Marie A, Coresh J, Folsom AR. HDL-cholesterol and the incidence of lung cancer in the Atherosclerosis Risk in Communities (ARIC) study. *Lung Cancer.* 2009; 61: 292–300.
9. Puchades-Carrasco L, Jantus-Lewintre E, Pérez-Rambla C, Camps C, Pineda-Lucena A. Serum metabolomic profiling facilitates the non-invasive identification of metabolic biomarkers associated with the onset and progression of non-small cell lung cancer. *Oncotarget.* 2016; 7: 12904–16. <https://doi.org/10.18632/oncotarget.7354> PMID: 26883203
10. Melkonian SC, Daniel CR, Ye Y, Pierzynski JA, Roth JA, Wu X. Glycemic Index, Glycemic Load, and Lung Cancer Risk in Non-Hispanic Whites. *Cancer Epidemiol Biomarkers Prev.* 2016; 25: 532–540. <https://doi.org/10.1158/1055-9965.EPI-15-0765> PMID: 26944871
11. Ho GYF, Zheng SL, Cushman M, Perez-soler R, Kim M, Xue X, et al. Associations of Insulin and IGFBP-3 with Lung Cancer Susceptibility in Current Smokers. *J Natl Cancer Inst.* 2016; 108: djw012. <https://doi.org/10.1093/jnci/djw012> PMID: 27071409
12. Freathy RM, Kazeem GR, Morris RW, Johnson PCD, Paternoster L, Ebrahim S, et al. Genetic variation at CHRNA5-CHRNA3-CHRNA4 interacts with smoking status to influence body mass index. *Int J Epidemiol.* 2011; 40: 1617–1628. <https://doi.org/10.1093/ije/dyr077> PMID: 21593077
13. Taylor AE, Morris RW, Fluharty ME, Bjorngaard JH, Åsvold BO, Gabrielsen ME, et al. Stratification by Smoking Status Reveals an Association of CHRNA5-A3-B4 Genotype with Body Mass Index in Never Smokers. *PLoS Genet.* 2014; 10: e1004799. <https://doi.org/10.1371/journal.pgen.1004799> PMID: 25474695
14. Winslow UC, Rode L, Nordestgaard BG. High tobacco consumption lowers body weight: a Mendelian randomization study of the Copenhagen General Population Study. *Int J Epidemiol.* 2015; 44: 540–550. <https://doi.org/10.1093/ije/dyu276> PMID: 25777141
15. Davey Smith G, Ebrahim S. Mendelian randomization: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003; 1–22. <https://doi.org/10.1093/ije/dyg070> PMID: 12689998
16. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet.* 2014; 23: 1–10.

17. Burgess S, Butterworth AS, Thompson JR. Beyond Mendelian randomization: how to interpret evidence of shared genetic predictors. *J Clin Epidemiol*. Elsevier; 2016; 69: 208–216. <https://doi.org/10.1016/j.jclinepi.2015.08.001> PMID: 26291580
18. Carreras-Torres R, Haycock PC, Relton CL, Martin RM, Smith GD, Kraft P, et al. The causal relevance of body mass index in different histological types of lung cancer: A Mendelian randomization study. *Sci Rep*. 2016; 6: 31121. <https://doi.org/10.1038/srep31121> PMID: 27487993
19. Gao C, Patel CJ, Michailidou K, Peters U, Gong J, Schildkraut J, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Int J Epidemiol*. 2016; dyw129. <https://doi.org/10.1093/ije/dyw129> PMID: 27427428
20. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013; 37: 658–65. <https://doi.org/10.1002/gepi.21758> PMID: 24114802
21. Burgess S, Scott R a, Timpson NJ, Davey-smith G, Thompson SG. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol*. Springer Netherlands; 2014; 1–39. <https://doi.org/10.1007/s10654-015-0011-z> PMID: 25773750
22. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. *Nature*. 2015; 526: 75–81. <https://doi.org/10.1038/nature15394> PMID: 26432246
23. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015; 518: 197–206. <https://doi.org/10.1038/nature14177> PMID: 25673413
24. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira Teresa, Locke AE, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*. 2015; 518: 187–196. <https://doi.org/10.1038/nature14132> PMID: 25673412
25. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013; 45: 1274–1283. <https://doi.org/10.1038/ng.2797> PMID: 24097068
26. Helgadóttir A, Gretarsdóttir S, Thorleifsson G, Hjartarson E, Sigurdsson A, Magnusdóttir A, et al. Variants with large effects on blood lipids and the role of cholesterol and triglycerides in coronary disease. *Nat Genet*. 2016; 48: 634–639. <https://doi.org/10.1038/ng.3561> PMID: 27135400
27. Scott RA, Lagou V, Welch RP, Wheeler E, May E. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet*. 2012; 44: 991–1005. <https://doi.org/10.1038/ng.2385> PMID: 22885924
28. McKay JD, Hun RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet*. 2017;
29. Timofeeva MN, Hung RJ, Rafnar T, Christiani DC, Field JK, Bickeböller H, et al. Influence of common genetic variation on lung cancer risk: Meta-analysis of 14 900 cases and 29 485 controls. *Hum Mol Genet*. 2012; 21: 4980–4995. <https://doi.org/10.1093/hmg/dds334> PMID: 22899653
30. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol*. 2014; 43: 922–929. <https://doi.org/10.1093/ije/dyu005> PMID: 24608958
31. Bowden J, Smith GD, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016; 40: 304–14. <https://doi.org/10.1002/gepi.21965> PMID: 27061298
32. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I² statistic. *Int J Epidemiol*. 2016; 1–14. <https://doi.org/10.1093/ije/dyw220>
33. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015; 512–525. <https://doi.org/10.1093/ije/dyv080> PMID: 26050253
34. Hemani G, Zheng J, Wade KH, Elsworth B, Langdon R, Burgess S, et al. MR-Base: a platform for systematic causal inference across the phenome using billions of genetic associations. *bioRxiv*. 2016; <https://doi.org/10.1101/078972>
35. Furberg H, Kim Y, Dackor J, Boerwinkle E, Franceschini N, Ardissono D, et al. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet*. 2010; 42: 441–447. <https://doi.org/10.1038/ng.571> PMID: 20418890

36. Esposito K, Chiodini P, Colao A, Lenzi A, Giugliano D. Metabolic Syndrome and Risk of Cancer. A systematic review and meta-analysis. *Diabetes Care*. 2012; 35: 2402–2411. <https://doi.org/10.2337/dc12-0336> PMID: [23093685](https://pubmed.ncbi.nlm.nih.gov/23093685/)
37. Radisauskas R, Kuzmickien I, Milinaviciene E, Everatt R. Hypertension, serum lipids and cancer risk: A review of epidemiological evidence. *Medicina (B Aires)*. 2016; 52: 89–98. <https://doi.org/10.1016/j.medici.2016.03.002> PMID: [27170481](https://pubmed.ncbi.nlm.nih.gov/27170481/)
38. Nead KT, Sharp SJ, Thompson DJ, Painter JN, Savage DB, Semple RK, et al. Evidence of a Causal Association Between Insulinemia and Endometrial Cancer: A Mendelian Randomization Analysis. *J Natl Cancer Inst*. 2015; 107: 1–7. <https://doi.org/10.1093/jnci/djv178> PMID: [26134033](https://pubmed.ncbi.nlm.nih.gov/26134033/)
39. Carreras-Torres R, Johansson M, Gaborieau V, Haycock PC, Wade KH, Relton CL, et al. The role of obesity, type 2 diabetes, and metabolic factors in pancreatic cancer: A Mendelian randomization study. *J Natl Cancer Inst*. 2017; 109: dxj012.
40. Braun S, Bitton-worms K, Leroith D. The Link between the Metabolic Syndrome and Cancer. *Int J Biol Sci*. 2011; 7: 1003–1015. PMID: [21912508](https://pubmed.ncbi.nlm.nih.gov/21912508/)
41. Kahn BB, Flier JS. Obesity and insulin resistance. *J Clin Invest*. 2000; 106: 473–481. <https://doi.org/10.1172/JCI10842> PMID: [10953022](https://pubmed.ncbi.nlm.nih.gov/10953022/)
42. Renehan AG, Zwahlen M, Egger M. Adiposity and cancer risk: new mechanistic insights from epidemiology. *Nat Rev Cancer*. 2015; 15: 484–498. <https://doi.org/10.1038/nrc3967> PMID: [26205341](https://pubmed.ncbi.nlm.nih.gov/26205341/)