



This is a repository copy of *Causal relationships between body mass index, smoking, and lung cancer: univariable and multivariable mendelian randomization.*

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/165485/>

Version: Accepted Version

---

**Article:**

Zhou, W., Liu, G., Hung, R.J. et al. (29 more authors) (2021) Causal relationships between body mass index, smoking, and lung cancer: univariable and multivariable mendelian randomization. *International Journal of Cancer*, 148 (5). pp. 1077-1086. ISSN 0020-7136

<https://doi.org/10.1002/ijc.33292>

---

This is the peer reviewed version of the following article: Zhou, W., Liu, G., Hung, R.J., Haycock, P.C., Aldrich, M.C., Andrew, A.S., Arnold, S.M., Bickeböller, H., Bojesen, S.E., Brennan, P., Brunnström, H., Melander, O., Caporaso, N.E., Landi, M.T., Chen, C., Goodman, G.E., Christiani, D.C., Cox, A., Field, J.K., Johansson, M., Kiemeny, L.A., Lam, S., Lazarus, P., Le Marchand, L., Rennert, G., Risch, A., Schabath, M.B., Shete, S.S., Tardón, A., Zienolddiny, S., Shen, H. and Amos, C.I. (2020), Causal Relationships between Body Mass Index, Smoking, and Lung Cancer: Univariable and Multivariable Mendelian Randomization. *Int. J. Cancer.*, which has been published in final form at <https://doi.org/10.1002/ijc.33292>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Causal Relationships between Body Mass Index, Smoking, and Lung Cancer: Univariable and Multivariable Mendelian Randomization

Wen Zhou<sup>1,2</sup>, Geoffrey Liu<sup>3</sup>, Rayjean J. Hung<sup>4,5</sup>, Philip C. Haycock<sup>6,7</sup>, Melinda C. Aldrich<sup>8</sup>, Angeline S. Andrew<sup>9</sup>, Susanne M. Arnold<sup>10</sup>, Heike Bickeböllner<sup>11</sup>, Stig E. Bojesen<sup>12</sup>, Paul Brennan<sup>13</sup>, Hans Brunnström<sup>14</sup>, Olle Melander<sup>14</sup>, Neil E. Caporaso<sup>15</sup>, Maria Teresa Landi<sup>15</sup>, Chu Chen<sup>16</sup>, Gary E. Goodman<sup>16</sup>, David C. Christiani<sup>17</sup>, Angela Cox<sup>18</sup>, John K. Field<sup>19</sup>, Mikael Johansson<sup>20</sup>, Lambertus A. Kiemeny<sup>21</sup>, Stephen Lam<sup>22</sup>, Philip Lazarus<sup>23</sup>, Loïc Le Marchand<sup>24</sup>, Gad Rennert<sup>25</sup>, Angela Risch<sup>26</sup>, Matthew B. Schabath<sup>27</sup>, Sanjay S. Shete<sup>28</sup>, Adonina Tardón<sup>29</sup>, Shanbeh Zienolddiny<sup>30</sup>, Hongbing Shen<sup>1\*</sup>, Christopher I. Amos<sup>2\*</sup>

\*Equal Contributors

<sup>1</sup>Department of Epidemiology, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, China

<sup>2</sup>Institute for Clinical and Translational Research, Baylor College of Medicine, Houston, TX, USA

<sup>3</sup>Princess Margaret Cancer Center, University Health Network, Toronto, ON, Canada

<sup>4</sup>Prosserman Centre for Population Health Research, Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, ON, Canada

<sup>5</sup>Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

<sup>6</sup>MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

<sup>7</sup>Bristol Medical School, Population Health Sciences, University of Bristol, Bristol, UK

<sup>8</sup>Department of Thoracic Surgery and Division of Epidemiology, Vanderbilt University Medical Center, Nashville, TN, USA

<sup>9</sup>Department of Epidemiology, Geisel School of Medicine, Dartmouth College, Hanover, NH, USA

<sup>10</sup>Markey Cancer Center, University of Kentucky, Lexington, KY, USA

- <sup>11</sup>Department of Genetic Epidemiology, University Medical Center, Georg-August-Universität Göttingen, Göttingen, Germany
- <sup>12</sup>Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark
- <sup>13</sup>Genetic Epidemiology Group, International Agency for Research on Cancer, Lyon, France
- <sup>14</sup>Clinical Sciences, Lund University, Lund, Sweden
- <sup>15</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA
- <sup>16</sup>Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center and Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA
- <sup>17</sup>Departments of Environmental Health and Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
- <sup>18</sup>Academic Unit of Clinical Oncology, University of Sheffield, Sheffield, UK
- <sup>19</sup>Department of Molecular and Clinical Cancer Medicine, Roy Castle Lung Cancer Research Programme, The University of Liverpool Cancer Research Centre, Liverpool, UK
- <sup>20</sup>Department of Radiation Sciences, Umeå University, Umea, Sweden
- <sup>21</sup>Department for Health Evidence, Radboud university medical center, Nijmegen, The Netherlands
- <sup>22</sup>Department of Integrative Oncology, British Columbia Cancer Agency, Vancouver, BC, Canada
- <sup>23</sup>Department of Pharmaceutical Sciences, College of Pharmacy and Pharmaceutical Sciences, Washington State University, Spokane, WA, USA
- <sup>24</sup>Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI, USA
- <sup>25</sup>Department of Community Medicine and Epidemiology, Carmel Medical Center and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology and Clalit National Cancer Control Center, Haifa, Israel
- <sup>26</sup>Cancer Center Cluster Salzburg at PLUS, Department of Molecular Biology, University of Salzburg, Heidelberg, Austria
- <sup>27</sup>Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, USA
- <sup>28</sup>Department of Biostatistics, Division of Basic Sciences, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
- <sup>29</sup>Faculty of Medicine, University of Oviedo and ISPA and CIBERESP, Oviedo, Spain
- <sup>30</sup>The National Institute of Occupational Health (STAMI), Oslo, Norway

**Correspondence to:** Christopher I. Amos, Institute for Clinical and Translational Research, Baylor College of Medicine, E-mail: [chrisa@bcm.edu](mailto:chrisa@bcm.edu) or Hongbing Shen, Department of Epidemiology, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, E-mail: [hbshen@njmu.edu.cn](mailto:hbshen@njmu.edu.cn)

**Short title:** Causal Relationships between BMI, Smoking, and Lung Cancer

**Key words:** Causal relationship, body mass index, smoking phenotypes, lung cancer, Mendelian randomization

**Abbreviations:**

BMI: body mass index

CI: confidence interval

$F_{TS}$ : two-sample conditional F-statistic

GIANT: Genetic Investigation of ANthropometric Traits

GWAS: Genome wide association study

GSCAN: GWAS and Sequencing Consortium of Alcohol and Nicotine use

ILCCO: International Lung Cancer Consortium

IV: instrumental variable

IVW: inverse-variance weighted

LD: linkage disequilibrium

MAF: minor allele frequency

MR: Mendelian randomization

MVMR: multivariable Mendelian randomization

NOME: NO Measurement Error

OR: odds ratio

SNP: single nucleotide polymorphism

TRICL: Transdisciplinary Research in Cancer of the Lung

UVMR: univariable Mendelian randomization.

**Article category:** Research Articles

### **Novelty and Impact**

Associations between body mass index (BMI) and several cancers are well established. However, the causal relationship of BMI on lung cancer is contradictory among studies. Here, using univariable and multivariable Mendelian randomization (MR), we confirmed an inverse direct effect of BMI on lung adenocarcinoma and a direct causal effect on small cell lung cancer, but did not observe direct effect of BMI on lung squamous cell carcinoma. The results highlight the histology-specific impact of BMI on lung carcinogenesis.

## Abstract

At the time of cancer diagnosis, body mass index (BMI) is inversely correlated with lung cancer risk, which may reflect reverse causality and confounding due to smoking behavior. We used two-sample univariable and multivariable Mendelian randomization (MR) to estimate causal relationships of BMI and smoking behaviors on lung cancer and histological subtypes, based on an aggregated genome wide association studies (GWASs) analysis of lung cancer in 29,266 cases and 56,450 controls. We observed a positive causal effect for high BMI on occurrence of small cell lung cancer (odds ratio (OR) = 1.60, 95% confidence interval (CI) = 1.24-2.06,  $P = 2.70 \times 10^{-4}$ ). After adjustment of smoking behaviors using multivariable Mendelian randomization (MVMR), a direct causal effect on small cell lung cancer ( $OR_{MVMR} = 1.28$ , 95% CI = 1.06-1.55,  $P_{MVMR} = 0.011$ ), and an inverse effect on lung adenocarcinoma ( $OR_{MVMR} = 0.86$ , 95% CI = 0.77-0.96,  $P_{MVMR} = 0.008$ ) were observed. A weak increased risk of lung squamous cell carcinoma was observed for higher BMI in univariable Mendelian randomization (UVMR) analysis ( $OR_{UVMR} = 1.19$ , 95% CI = 1.01-1.40,  $P_{UVMR} = 0.036$ ), but this effect disappeared after adjustment of smoking ( $OR_{MVMR} = 1.02$ , 95% CI = 0.90-1.16,  $P_{MVMR} = 0.746$ ). These results highlight the histology-specific impact of BMI on lung carcinogenesis and imply mediator role of smoking behaviors in the association between BMI and lung cancer.

## Introduction

Although body mass index (BMI) is known to be associated with the risk of several cancers<sup>1</sup>, many epidemiological studies<sup>1-6</sup> have found that BMI is inversely correlated with lung cancer risk. At the time of or 1 to 2 year leading up to cancer diagnosis, lung cancer cases tend to have lower BMI compared to age, population, and sex matched controls<sup>7</sup>. This observation may reflect reverse causality related to the latent effects of lung cancer on BMI due to cachexia and the disease process<sup>8</sup> and confounding due to smoking behavior<sup>4, 9-11</sup>, which suppresses appetite.

Mendelian randomization (MR) complements traditional epidemiological methods as it uses genetic variants as instrumental variables (IVs) to estimate causal effects and is free from the effects of confounders and reverse causality<sup>12</sup> but could be influenced by pleiotropy or linkage disequilibrium among the markers<sup>13</sup>. The conventional, univariable MR analysis evaluates the effect of a single predictor on an outcome. There are three assumptions for a valid IV – it must be: (i) associated with the exposure (the “relevance” assumption); (ii) independent of the outcome given the exposure (the “exclusion restriction”); and (iii) independent of all (both observed and unobserved) confounders (the “exchangeability” assumption).<sup>14, 15</sup> If an IV is associated with a confounder of exposure and outcome then there is a conflict with these assumptions, which may lead to potential biases and erroneous conclusions. If associations between the IV and both the exposure and the confounder are measured, then their effects on the outcome can be estimated jointly using multivariable MR<sup>16</sup>.

Multivariable MR is an extension of univariable MR that takes pleiotropy among multiple traits into account. The assumptions of multivariable MR is broader than that of univariable MR: genetic variants could influence multiple measured exposures, and the exclusion restriction and exchangeability assumption extend accordingly.<sup>16</sup> Under the scenario that there is a secondary exposure acts as a mediator of the relationship between primary exposure and outcome, multivariable MR can provide a consistent estimator of the direct effect of the exposure on the outcome that does not work via the mediator.<sup>17</sup>

In this study, we performed Mendelian randomization analyses to investigate causal effects of BMI and smoking phenotypes on lung cancer using univariable MR and multivariable MR methods. Sensitivity analyses were conducted to evaluate the impact that assumptions had on the findings and to ensure the robustness of the results.

## Materials and Methods

### *Genetic instruments for BMI and smoking phenotypes*

Genetic instruments for BMI were identified using results from the largest available meta-analysis of genome wide association studies (GWASs) for BMI, which combined summary statistics from the Genetic Investigation of ANthropometric Traits (GIANT) consortium<sup>18</sup> with GWAS for BMI performed in UK Biobank participants of European ancestry, reaching about 700,000 individuals.<sup>19</sup> The former association testing was performed with the inverse normally transformed BMI residuals using linear regression assuming an additive genetic model. The associations identified in the latter study were estimated using linear mixed model assuming an infinitesimal model. METAL<sup>20</sup> was used to perform fixed-effect inverse variance weighted meta-analysis.

Genetic instruments for smoking phenotypes were obtained from the most recent meta-analysis on tobacco and alcohol consumption based on over 30 GWASs, which identified 566 genetic variants associated with four smoking phenotypes in over 1.2 million individuals of European ancestry.<sup>21</sup> Detailed information about the smoking phenotypes have been described elsewhere<sup>21</sup>. Briefly, they included smoking status (Ever/Never), age of smoking initiation, smoking cessation (Current/Former), and number of cigarettes smoked per day (either as a current smoker or former smoker, quantitative measures were binned into five bins or collected with pre-defined bins as follows: 1 = 1-5, 2 = 6-15, 3 = 16-25, 4 = 26-35, 5 = 36+ cigarettes per day). The association statistics were generated using a linear mixed model for all phenotypes in family studies and quasi-continuous phenotypes (age of smoking initiation and cigarettes per day) in unrelated individuals. A logistic model was utilized to estimate additive genetic effects for binary phenotypes (smoking status and smoking cessation) in studies of unrelated individuals. Meta-analysis was performed using rareGWAMA<sup>22</sup>.

The IVs for MR analyses were selected based on the following criteria: (i)  $r^2$  measure of linkage disequilibrium (LD) among instruments  $< 0.01$  at a 10Mb window; (ii)  $P$  value less than the genome-wide significant level identified in the corresponding study ( $1 \times 10^{-8}$  for BMI,  $5 \times 10^{-8}$  for smoking phenotypes); (iii) minor allele frequency (MAF)  $> 0.01$ ; (iv) non-palindromic single nucleotide polymorphisms (SNPs) (A/T and G/C polymorphisms were considered for removal but there was no ambiguity in effects). We further split the genetic instruments for BMI into those that only affect BMI ("BMI-Only SNP") and those that affect both smoking and BMI ("BMI&Smoking SNP") by using the  $P$  values in the tobacco use GWAS meta-analysis: If the  $P$  value of any of smoking phenotypes was less than 0.05, it was classified as the "BMI&Smoking SNP".



### *Genetic associations of SNPs with lung cancer risk*

Summary statistics on lung cancer risk, including odds ratio (OR) estimates and standard errors for instrumental SNPs, were available from the Transdisciplinary Research in Cancer of the Lung (TRICL) and International Lung Cancer Consortium (ILCCO) based on an aggregated GWAS analysis of lung cancer in 29,266 cases and 56,450 controls.<sup>23</sup> The associations between instrumental SNPs and histological subtypes of lung cancer were also provided in the original study. These analyses were restricted to European-descent participants, defined as having 80% or more derived European ancestry.

### *Mendelian Randomization*

The relationship between BMI, smoking, and lung cancer is illustrated in Figure 1. The direct effect of BMI on lung cancer is the effect that not via smoking, which is equal to  $\beta_{XZ}$ . The total effect of BMI on lung cancer is the effect of BMI on lung cancer directly plus the effect of BMI on lung cancer via smoking, which is equal to  $\beta_{XZ} + \alpha_{XY}\beta_{YZ}$ . We performed univariable MR analyses for BMI and smoking on lung cancer risk separately to estimate the total causal effects of BMI on lung cancer and smoking behaviors on lung cancer. To evaluate whether BMI was causal for smoking phenotypes or smoking phenotypes have causal effects on BMI, we conducted bidirectional MR for BMI and smoking phenotypes. Specifically, MR analyses were first performed from BMI to smoking phenotypes, and then the direction was reversed.

To investigate the direct effects of BMI and smoking behaviors on lung cancer, we performed multivariable MR<sup>16</sup> analysis, which is an extension of univariable MR that allows detecting causal effects of multiple risk factors jointly.<sup>17</sup> Multivariable MR takes into account the relationship between BMI and smoking phenotypes and the fact that the SNPs used in the MR analyses are often associated with both BMI and smoking (i.e. SNPs in our analyses were associated with BMI and at least one smoking phenotype with  $P < 0.05$ , the “BMI&Smoking SNP”). The SNPs used to conduct multivariable MR were combinations of instrumental variables of each exposure. We restricted the analysis to SNPs that were clumped on  $r^2 < 0.01$  within 10Mb. MR analyses and clumping were conducted using the packages “TwoSampleMR”<sup>24</sup> and “MendelianRandomization” (<https://cran.r-project.org/package=MendelianRandomization>) in R.

## *Sensitivity analysis*

We applied several MR methods to perform MR analyses. For univariable MR, we used inverse-variance weighted (IVW)<sup>25, 26</sup>, weighted median<sup>27</sup>, weighted mode<sup>28</sup>, and MR-Egger<sup>13</sup> approaches. For multivariable MR, we used regression-based IVW<sup>16</sup> and Egger<sup>29</sup>. We calculated  $R^2$  to estimate the proportion of phenotypic variance explained and F-statistics to evaluate the strength of instruments<sup>30</sup>. The two-sample conditional F-statistics ( $F_{TS}$ )<sup>17, 31</sup> were also estimated to evaluate the strength of SNP-exposure conditional on other exposures. The intercept test of MR-Egger was used to assess horizontal pleiotropic effects,  $I_G \chi^2$  statistic to check the violation of “NOME (NO Measurement Error) assumption”, which states the potential relative bias due to measurement error. The heterogeneity estimated by Cochran’s Q test was to appraise whether any single instrumental variable was driving the results and to check for consistency of the analyses with MR assumptions. The outliers in IVW and MR-Egger regression models was identified using RadialMR<sup>32</sup>. In addition, we calculated the power of our study using the mRnd power calculator for conventional MR.<sup>33</sup>

## **Results**

### *Genetic instruments*

In total, 842 independent SNPs were included as instrumental variables for BMI and 110, 7, 10, and 29 SNPs for smoking phenotypes of smoking status, age of smoking initiation, smoking cessation, and cigarettes per day, respectively (Supporting Information Table S1). The F-statistic values (Supporting Information Table S2) for individual SNPs ranged from 30 to 1360, with means of 64, 41, 39, 53, and 92 for BMI, smoking status, age of smoking initiation, smoking cessation, and cigarettes per day, respectively. However, the two-sample conditional F-statistics were all less than 10 for BMI and all smoking phenotypes, indicating the strength of genetic instruments was significantly lower after conditioning on other exposures, which may result in greatly reduced power for multivariable MR to estimate causal effects. Despite a reduction in strength, the BMI instrument had an  $F_{TS}$  over 8 reflecting moderate strength<sup>31</sup>. The variability explained by these genetic instruments were 7.33% for BMI, 0.71% for smoking status, 0.1% for age of smoking initiation, 0.17% for smoking cessation, and 1.03% for cigarettes per day (Supporting Information Table S2). Power calculations for the univariable IVW MR analyses (Supporting Information Figure S1) indicated greater than 80% statistical power to detect an odds ratio bigger than 1.077 for per 1 unit

increase of BMI, 1.26 for ever smoking, 1.77 for one age younger of smoking initiation, 1.57 for current smoking, and 1.213 for per 10 cigarettes increase of cigarettes per day. There were 451 instrumental variables for BMI identified as “BMI&Smoking SNP” and 390 identified as “BMI-Only SNP” (Supporting Information Table S1).

### *Univariable MR analyses*

From the univariable MR analyses (Table 1, Supporting Information Table S3 and Figure S2) we observed that increased BMI has a positive effect on occurrence of overall lung cancer when using all BMI associated SNPs (OR = 1.19, 95% confidence interval (CI) = 1.10-1.28,  $P = 7.24 \times 10^{-6}$ ). This effect was also significant in histological subtypes of lung squamous cell carcinoma (OR = 1.36, 95% CI = 1.22-1.52,  $P = 3.59 \times 10^{-8}$ ) and small cell lung cancer (OR = 1.73, 95% CI = 1.47-2.04,  $P = 5.50 \times 10^{-11}$ ) but not in lung adenocarcinoma (OR = 1.00, 95% CI = 0.91-1.10,  $P = 0.935$ ) (the first four rows of Table 1). The causal relationships estimated by BMI&Smoking SNP were consistent with the all SNPs set, with a slightly stronger effect (the 5-8 rows of Table 1). The BMI-Only SNP IVs were significant for small cell lung cancer (OR = 1.60, 95% CI = 1.24-2.06,  $P = 2.70 \times 10^{-4}$ ); the associations between BMI and other lung cancer subtypes were weak or nonexistent (the last four rows of Table 1). As smoking behaviors are confounders of BMI and lung cancer, the BMI-Only SNPs, which have excluded genetic variants that have significant association with smoking phenotypes, should be instruments that are more valid for BMI in univariable MR analysis. The univariable MR using BMI-Only SNP estimates the total effect of BMI on lung cancer.<sup>17</sup>

Among the four smoking phenotypes (Supporting Information Table S4 and Figure S3), smoking status was positively associated with lung cancer and histological subtypes, which is consistent with known associations of smoking with lung cancer risk. Age of smoking initiation was inversely associated with lung squamous cell carcinoma and small cell lung cancer, which indicated that the earlier an individual started regular smoking, the greater his/her risk of lung squamous cell carcinoma and small cell lung cancer. There was an expected strong evidence of causal effects of cigarettes per day (per 10 cigarettes increase) with lung cancer and histological subtypes (Overall: OR = 2.89, 95% CI = 2.33-3.58,  $P = 3.59 \times 10^{-22}$ ; Adenocarcinoma: OR = 2.57, 95% CI = 2.06-3.20,  $P = 6.20 \times 10^{-17}$ ; Squamous cell: OR = 3.27, 95% CI = 2.60-4.10,  $P = 1.66 \times 10^{-24}$ ; Small cell: OR = 3.32, 95% CI = 2.52-4.37,  $P = 1.80 \times 10^{-17}$ ). No significant causal effects of the genetic instrument modeling smoking cessation on lung cancer was observed. The sensitivity analyses using median weighted IVs was similar for all exposures (ORs were comparable to those estimated by the IVW method), except for

smoking cessation. Results for MR Egger regression were also qualitatively similar to results estimated by the IVW method, except smoking cessation for which protective effects were found (Supporting Information Table S4 and Figure S3). These results suggest heterogeneity in instrumental effects for the smoking cessation variable but not for other variables.

The MR analyses for evaluating the effect of BMI on smoking phenotypes yielded consistent results across three SNP sets (Supporting Information Table S5 and Figure S4). Increased BMI has positive effects on smoking status (OR = 1.24, 95% CI = 1.19-1.28,  $P = 6.60 \times 10^{-36}$ ), smoking cessation (OR = 1.14, 95% CI = 1.11-1.18,  $P = 7.20 \times 10^{-15}$ ), and cigarettes per day (OR = 1.39, 95% CI = 1.34-1.44,  $P = 7.54 \times 10^{-64}$ ), while inverse effect on age of smoking initiation (OR = 0.91, 95% CI = 0.90-0.93,  $P = 3.68 \times 10^{-26}$ ). There was limited evidence (Supporting Information Table S5 and Figure S5) for effects of smoking behaviors on BMI. In addition, no significant horizontal pleiotropic effects were detected in MR Egger (for the intercept of MR Egger, all  $P$  value more than 0.1). The low  $I_G^2$  statistics (less than 80%) for all smoking phenotypes suggest bias due to measurement errors in MR Egger (the last four rows of Supporting Information Table S5).

### *Multivariable MR analyses*

We estimated mutually the effects of smoking and BMI on lung cancer using multivariable MR and observed a directly inverse effect of BMI on lung adenocarcinoma (OR = 0.86, 95% CI = 0.77-0.96,  $P = 0.008$ ) and a weak risk effect on small cell lung cancer (OR = 1.28, 95% CI = 1.06-1.55,  $P = 0.011$ ) (Table 2 and Figure 2). After adjustment of BMI and other smoking phenotypes, smoking status still had a direct effect on overall lung cancer, lung adenocarcinoma and lung squamous cell carcinoma, but not on small cell lung cancer. There were positive direct effects of smoking cessation on overall lung cancer, squamous and small cell lung cancer, but no direct effect on lung adenocarcinoma. Cigarettes per day has direct effects on overall lung cancer, squamous and small cell lung cancer. No significant direct effect was detected for age of smoking initiation on lung cancer risk, while jointly modeling BMI and other smoking phenotypes.

We also investigated the direct effects between BMI and smoking behaviors using multivariable MR (Supporting Information Table S6 and Table S7). Strong evidence for direct effects of BMI on smoking status and cigarettes per day were observed, as well as inverse effect on age of smoking initiation. There was also some evidence for direct

effects of smoking behaviors on BMI, such as smoking status, smoking cessation and cigarettes per day.

### *Sensitivity analysis*

Since genome-wide significant level for each exposure was determined according to the corresponding study of its genetic instruments<sup>19, 21</sup>, the thresholds used for BMI ( $1 \times 10^{-8}$ ) and smoking phenotypes ( $5 \times 10^{-8}$ ) were different. To provide a sensitivity analysis, we selected IVs based on instruments satisfying either  $5 \times 10^{-8}$  or  $1 \times 10^{-8}$  and performed additional MR analyses. As shown in Supporting Information Table S8-S10, the causal effects of BMI and smoking behaviors on lung cancer were robust using different thresholds.

In univariable MR analyses, the  $I_G^2$  statistics of MR Egger for BMI and smoking phenotypes on lung cancer (Supporting Information Table S3 and Table S4) were less than 90%, except that of cigarettes per day on lung cancer and its subtypes (range from 96.2% to 96.5%). The lower values of  $I_G^2$  indicates violation of the “NOME” assumption in the causal estimates due to variations in the genetic associations. In addition, the Cochran’s Q test in the IVW model and MR Egger model (Supporting Information Table S11) suggested that there was strong evidence of heterogeneity in most of the instrumental variables. However, we did not observe significant directional pleiotropy using MR Egger, as the  $P$  value for the intercept tests were not significant (Supporting Information Table S3 and Table S4), which suggests balanced pleiotropy<sup>13</sup> in the genetic instruments of univariable MR analyses. Balanced pleiotropy means that the pleiotropic effects of genetic instruments are balanced around the overall effect. The symmetry in the funnel plots also supported balanced pleiotropy (Supporting Information Figure S6-S7). However, outliers in the scatter plot suggested presence of heterogeneity. (Supporting Information Figure S8-S9). The outliers in the analysis of univariable MR for smoking cessation on lung cancer risk were rs518425 (intronic of *CHRNA5*) and rs56113850 (intronic of *CYP2A6*). The effects of these two large-effect SNPs appear to be opposite for BMI phenotypes, which explains the symmetry. After removing these two SNPs, no heterogeneity for smoking cessation remained and the causal relationship between smoking cessation and lung cancer was significant (Supporting Information Table S12 and Figure S10). In addition, the estimates of MR IVW, MR Weighted Median, MR Weighted Mode, and MR Egger show nearly consistent effects except for smoking cessation, which has strong evidence of heterogeneity (Supporting Information Figure S2-S5).

The estimates were robust between multivariable MR IVW and multivariable MR Egger analyses, which suggested reliable evaluation of causal effects for BMI and smoking phenotypes on lung cancer risk (Table 2, Figure 2 and Supporting Information Figure S11). In addition, there was no significant evidence for a non-zero intercept of multivariable MR Egger regression, which also supports the reliability of the results for multivariable MR analyses.

### *MR analyses for lung cancer among ever-/never-smokers*

To investigate the role of smoking phenotypes in the causal relationship between BMI and lung cancer, we further performed MR analyses for BMI and smoking phenotypes on ever- versus never-smokers for lung cancer. A positive effect of high BMI on occurrence of lung cancer among ever-smokers was observed using BMI-All SNPs or BMI&Smoking SNPs. An inverse effect of BMI on risk for lung cancer among never-smokers was identified using BMI-Only SNPs (Supporting Information Table S13). After adjustment of smoking phenotypes using multivariable MR, the estimated causal effects of BMI on ever- or never-smokers lung cancer was further attenuated (Supporting Information Table S14). A weakly but not significant protective effect of BMI on lung cancer risk among never smokers was observed, but further much larger studies are needed to evaluate these associations.

## **Discussion**

In this study, we observed that adjusted for smoking, BMI decreased risk of lung adenocarcinoma and increased risk of small cell lung cancer. BMI increased risk of lung squamous cell carcinoma but this effect was mediated by smoking. This study also highlights the effectiveness and necessity of multivariable MR in MR analyses, especially when large amounts of genetic variants are used as instrumental variables.

Some previous observational studies<sup>5, 6, 34-37</sup> have demonstrated that BMI was associated with decreasing risk of lung cancer, even after adjustment for smoking and other confounders<sup>5</sup>, but separating out residual confounding with smoking is difficult in observational analyses. The inverse association for lung adenocarcinoma and positive association for small cell carcinoma were also observed in a prospective cohort study<sup>6</sup>. There was no heterogeneity between males and females, but the association was stronger in ever-smokers than in never-smokers<sup>6, 34, 35</sup>. In our findings, there was evidence of a causal effect for overall lung cancer among ever-smokers if using the SNPs associated with both BMI and smoking behavior, while reverse effect for overall lung cancer among never-smokers if using the SNPs only associated with BMI,

suggests there might be some interactions between smoking phenotypes and BMI in the process of impacting lung carcinogenesis.

Carreras-Torres et al. found that increased BMI, as defined by instruments for obesity from Mendelian Randomization, was associated with increased risks for lung squamous cell and small cell lung cancers but not for lung adenocarcinomas.<sup>38</sup> These were consistent with our univariable MR analyses using SNPs only associated with BMI, which indicated a total effect for BMI on lung cancer.<sup>17</sup> Moreover, our multivariable MR analyses also identified direct effects for BMI with the risk of lung adenocarcinomas and small cell lung cancers. Other Mendelian randomization studies for BMI and lung cancer also identified evidence of an increased risk for overall lung cancer<sup>39</sup>, lung squamous cell carcinoma<sup>39, 40</sup>, and small cell lung cancer<sup>40</sup>, as well as evidence of a decreased risk for lung adenocarcinoma<sup>41</sup>. However, these analyses have not jointly considered modeling smoking with BMI, and we have seen that there are complex relationships among these factors, such that both need to be jointly modeled.

Smoking behavior is a confounder when evaluating the association between BMI and lung cancer. Previous MR study confirmed the role of obesity in smoking behaviors.<sup>9</sup> By using multivariable MR, we removed bias caused by smoking phenotypes compared to using univariable MR. In the bidirectional analysis, we found that there was a much stronger path from BMI to smoking phenotypes rather than from smoking phenotypes to BMI, which indicates the role of smoking as a potential mediator between BMI and lung cancer risk. On the other hand, the effects of smoking (such as smoking cessation and cigarettes per day) on BMI were also observed. These results implied complicated effects of BMI and smoking behaviors on lung cancer risk. Further analyses such as network Mendelian randomization, which uses genetic instruments to investigate mediation in causal pathways, may provide insights into the causal relationships.<sup>42</sup>

The possible mechanism between obesity and lung cancer has been investigated in previous studies.<sup>43</sup> Adipokines, which are secreted by adipose tissue, have properties affecting carcinogenesis, immunomodulation, appetite and energy homeostasis regulation, and variations in levels have been associated with tumor susceptibility and pathogenesis<sup>44, 45</sup>. In particular, leptin seems to mediate and maintain chronic inflammation after exposure to inhaled antigens (e.g., smoke particles) and induce lung carcinogenesis<sup>46</sup>. In addition, after adding recombinant human leptin, cell proliferation was observed in SQ-5 human clonal squamous lung cancer derived cell lines<sup>47</sup>. Adiponectin, unlike leptin, correlates inversely with body weight and is considered a protective hormone that exerts anti-inflammatory effects. Petridou et al.<sup>48</sup> demonstrated the expression of adiponectin receptors exclusively in lung cancer tissues, which

supports the hypothesis that adiponectin mediates lung cancer development. Nevertheless, the role of obesity in lung carcinogenesis remains unclear and inconclusive. As obesity may influence different histological subtypes of lung cancer, further research is warranted to elucidate the role of obesity in lung cancer.

Although two sample Mendelian Randomization is a powerful approach to make causal inference between exposures and outcome using summary statistics, we should be prudent with our findings because of several limitations. First, sample overlap can result in inflation of test results,<sup>49</sup> but we expect little effect since there are no known overlaps in samples. In addition, the sample size for histological subtypes of lung cancer were limited, which could lead to false negative errors (lung adenocarcinoma: 11,273 cases and 55,483 controls; lung squamous cell carcinoma: 7,426 cases and 55,627 controls; small cell lung cancer: 2,664 cases and 21,444 controls; ever-smokers lung cancer: 23,223 cases and 16,964 controls; never-smokers lung cancer: 2,355 cases and 7,504 controls<sup>23</sup>). Moreover, according to previous studies<sup>50, 51</sup>, in the situation of both exposure and outcome are binary, Wald-type estimators could introduce bias for causal odds ratio. The causal estimations between binary smoking phenotypes (such as smoking status and smoking cessation) and lung cancer should be evaluated carefully. Nevertheless, we believe that the impact is small since the estimation will be close to the true value if the sample size is large,<sup>15</sup> and we also had a large number of instrumental variables. Furthermore, the BMI changes with age and the MR analysis was designed to estimate the lifetime effects of exposure<sup>52</sup>, but the genetic instruments used here were associated with BMI at a specific age interval, so it may impact interpretability and validity of the results.

Taken together, we demonstrate that high BMI has direct inverse association with risk of lung adenocarcinoma and positive causal direct effects with small cell lung cancer. The total effects of BMI on lung squamous cell carcinoma and small cell lung cancer may be mediated by smoking phenotypes. These findings reveal heterogeneity of causal effects of BMI on histological subtypes of lung cancer. We expect that further investigations are needed to demystify the complex causal relationships between BMI, smoking and lung cancer. Our findings also suggest that future studies could explore the relevance of change BMI for evaluation risk of lung cancer.

## **Acknowledgements**

Dr. Amos is a research scholar of the Cancer Prevention Research Institute of Texas (CPRIT). His research is partially funded by CPRIT grant RR170048 and by NIH/NCI



grant U19CA203654. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization. The Boston Lung Cancer Study was funded by NIH (NCI) U01CA209414 (PI: Christiani). The lung eQTL study at Laval University was supported by the Fondation de l'Institut universitaire de cardiologie et de pneumologie de Québec and the Canadian Institutes of Health Research (MOP – 123369). The EAGLE study was supported by the Intramural Research Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS. The Multiethnic Cohort Study is supported by National Institutes of Health (CA164973). The CARET study was supported by the National Institutes of Health/National Cancer Institute: UM1 CA167462 (PI: Goodman), U01CA6367307 (PIs: Omen, Goodman); R01 CA111703 (PI: Chen), 5R01 CA151989-01A1 (PI: Doherty) and U01 CA167462 (PI: Chen). Philip C Haycock is supported by Cancer Research UK (C18281/A19169). Wen Zhou was supported by China Scholarship Council and Nanjing Medical University.

## **Conflict of interests**

The authors declare no competing interests.

## **Data accessibility**

GWAS summary statistics for BMI can be downloaded from the GIANT consortium website ([https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT\\_consortium\\_data\\_files](https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files)) and from the Program in Complex Trait Genomics website (<https://cnsgenomics.com/content/data>). GWAS summary statistics for tobacco use can be downloaded from the GWAS and Sequencing Consortium of Alcohol and Nicotine use (GSCAN) consortium website (<https://genome.psych.umn.edu/index.php/GSCAN>). Other data that support the findings of our study are available on request from the corresponding author.

## **Ethics statement**

Each study has been approved by the Ethical Committees of the original studies.

## References

1. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5-24 million UK adults. *The Lancet* 2014;**384**: 755-65.
2. Yang Y, Dong J, Sun K, Zhao L, Zhao F, Wang L, Jiao Y. Obesity and incidence of lung cancer: a meta-analysis. *Int J Cancer* 2013;**132**: 1162-9.
3. Song X, Pukkala E, Dyba T, Tuomilehto J, Moltchanov V, Mannisto S, Jousilahti P, Qiao Q. Body mass index and cancer incidence: the FINRISK study. *Eur J Epidemiol* 2014;**29**: 477-87.
4. Smith L, Brinton LA, Spitz MR, Lam TK, Park Y, Hollenbeck AR, Freedman ND, Gierach GL. Body mass index and risk of lung cancer among never, former, and current smokers. *J Natl Cancer Inst* 2012;**104**: 778-89.
5. Dewi NU, Boshuizen HC, Johansson M, Vineis P, Kampman E, Steffen A, Tjonneland A, Halkjaer J, Overvad K, Severi G, Fagherazzi G, Boutron-Ruault MC, et al. Anthropometry and the Risk of Lung Cancer in EPIC. *Am J Epidemiol* 2016;**184**: 129-39.
6. Yu D, Zheng W, Johansson M, Lan Q, Park Y, White E, Matthews CE, Sawada N, Gao YT, Robien K, Sinha R, Langhammer A, et al. Overall and Central Obesity and Risk of Lung Cancer: A Pooled Analysis. *J Natl Cancer Inst* 2018;**110**: 831-42.
7. El-Zein M, Parent ME, Nicolau B, Koushik A, Siemiatycki J, Rousseau MC. Body mass index, lifetime smoking intensity and lung cancer risk. *Int J Cancer* 2013;**133**: 1721-31.
8. Schmitt R. Quality of Life Issues in Lung Cancer. *Chest* 1993;**103**: 51S-5S.
9. Carreras-Torres R, Johansson M, Haycock PC, Relton CL, Davey Smith G, Brennan P, Martin RM. Role of obesity in smoking behaviour: Mendelian randomisation study in UK Biobank. *BMJ* 2018;**361**: k1767.
10. 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**: 540-50.
11. Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. *Eur Respir J* 2016;**48**: 889-902.
12. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;**32**: 1-22.
13. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;**44**: 512-25.
14. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res* 2007;**16**: 309-30.
15. Didelez V, Meng S, Sheehan NA. Assumptions of IV Methods for Observational Epidemiology. *Statistical Science* 2010;**25**: 22-40.
16. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* 2015;**181**: 251-60.
17. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol* 2019;**48**: 713-27.
18. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, Powell C, Vedantam S, Buchkovich ML, Yang J, Croteau-Chonka DC, Esko T, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;**518**: 197-206.
19. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, Frayling TM, Hirschhorn J, Yang J, Visscher PM, Consortium G. Meta-analysis of genome-wide association studies for height and

- body mass index in approximately 700000 individuals of European ancestry. *Hum Mol Genet* 2018;**27**: 3641-9.
20. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;**26**: 2190-1.
  21. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, Datta G, Davila-Velderrain J, McGuire D, Tian C, Zhan X, and Me Research T, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet* 2019;**51**: 237-44.
  22. Liu DJ, Peloso GM, Zhan X, Holmen OL, Zawistowski M, Feng S, Nikpay M, Auer PL, Goel A, Zhang H, Peters U, Farrall M, et al. Meta-analysis of gene-level tests for rare variant association. *Nat Genet* 2014;**46**: 200-4.
  23. McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, Caporaso NE, Johansson M, Xiao X, Li Y, Byun J, Dunning A, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet* 2017;**49**: 1126-32.
  24. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, Tan VY, Yarmolinsky J, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018;**7**.
  25. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013;**37**: 658-65.
  26. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG, Consortium E-I. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol* 2015;**30**: 543-52.
  27. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* 2016;**40**: 304-14.
  28. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* 2017;**46**: 1985-98.
  29. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat Med* 2017;**36**: 4705-18.
  30. Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, Davey Smith G, Sterne JA. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res* 2012;**21**: 223-42.
  31. Sanderson E, Spiller W, Bowden J. Testing and Correcting for Weak and Pleiotropic Instruments in Two-Sample Multivariable Mendelian Randomisation. *bioRxiv* 2020: 2020.04.02.021980.
  32. Bowden J, Spiller W, Del Greco MF, Sheehan N, Thompson J, Minelli C, Davey Smith G. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. *Int J Epidemiol* 2018;**47**: 1264-78.
  33. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol* 2013;**42**: 1497-501.
  34. Everatt R, Virviciute D, Kuzmickiene I, Tamosiunas A. Body mass index, cholesterol level and risk of lung cancer in Lithuanian men. *Lung Cancer* 2014;**85**: 361-5.
  35. Duan P, Hu C, Quan C, Yi X, Zhou W, Yuan M, Yu T, Kourouma A, Yang K. Body mass index and risk of lung cancer: Systematic review and dose-response meta-analysis. *Sci Rep* 2015;**5**: 16938.
  36. Benn M, Tybjaerg-Hansen A, Smith GD, Nordestgaard BG. High body mass index and cancer risk—a Mendelian randomisation study. *Eur J Epidemiol* 2016;**31**: 879-92.

37. Sanikini H, Yuan JM, Butler LM, Koh WP, Gao YT, Steffen A, Johansson M, Vineis P, Goodman GE, Barnett MJ, Hung RJ, Chen C, et al. Body mass index and lung cancer risk: a pooled analysis based on nested case-control studies from four cohort studies. *BMC Cancer* 2018;**18**: 220.
38. Carreras-Torres R, Haycock PC, Relton CL, Martin RM, Smith GD, Kraft P, Gao C, Tworoger S, Le Marchand L, Wilkens LR, Park SL, Haiman C, 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.
39. Gao C, Patel CJ, Michailidou K, Peters U, Gong J, Schildkraut J, Schumacher FR, Zheng W, Boffetta P, Stucker I, Willett W, Gruber S, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Int J Epidemiol* 2016;**45**: 896-908.
40. Carreras-Torres R, Johansson M, Haycock PC, Wade KH, Relton CL, Martin RM, Davey Smith G, Albanes D, Aldrich MC, Andrew A, Arnold SM, Bickeboller H, et al. Obesity, metabolic factors and risk of different histological types of lung cancer: A Mendelian randomization study. *PLoS One* 2017;**12**: e0177875.
41. Brennan P, McKay J, Moore L, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, et al. Obesity and cancer: Mendelian randomization approach utilizing the FTO genotype. *Int J Epidemiol* 2009;**38**: 971-5.
42. Burgess S, Daniel RM, Butterworth AS, Thompson SG, Consortium EP-I. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. *Int J Epidemiol* 2015;**44**: 484-95.
43. Ntikoudi E, Kiagia M, Boura P, Syrigos KN. Hormones of adipose tissue and their biologic role in lung cancer. *Cancer Treat Rev* 2014;**40**: 22-30.
44. Chen J. Multiple signal pathways in obesity-associated cancer. *Obes Rev* 2011;**12**: 1063-70.
45. Khandekar MJ, Cohen P, Spiegelman BM. Molecular mechanisms of cancer development in obesity. *Nat Rev Cancer* 2011;**11**: 886-95.
46. Ribeiro R, Araújo A, Lopes C, Medeiros R. Immunoinflammatory Mechanisms in Lung Cancer Development: Is Leptin a Mediator? *Journal of Thoracic Oncology* 2007;**2**: 105-8.
47. Tsuchiya T, Shimizu H, Horie T, Mori M. Expression of leptin receptor in lung: leptin as a growth factor. *Eur J Pharmacol* 1999;**365**: 273-9.
48. Petridou ET, Mitsiades N, Gialamas S, Angelopoulos M, Skalkidou A, Dessypris N, Hsi A, Lazaris N, Polyzos A, Syrigos C, Brennan AM, Tseleni-Balafouta S, et al. Circulating adiponectin levels and expression of adiponectin receptors in relation to lung cancer: two case-control studies. *Oncology* 2007;**73**: 261-9.
49. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol* 2016;**40**: 597-608.
50. Palmer TM, Sterne JA, Harbord RM, Lawlor DA, Sheehan NA, Meng S, Granell R, Smith GD, Didelez V. Instrumental variable estimation of causal risk ratios and causal odds ratios in Mendelian randomization analyses. *Am J Epidemiol* 2011;**173**: 1392-403.
51. Disney-Hogg L, Cornish AJ, Sud A, Law PJ, Kinnersley B, Jacobs DI, Ostrom QT, Labreche K, Eckel-Passow JE, Armstrong GN, Claus EB, Il'yasova D, et al. Impact of atopy on risk of glioma: a Mendelian randomisation study. *BMC Med* 2018;**16**: 42.
52. Labrecque JA, Swanson SA. Interpretation and Potential Biases of Mendelian Randomization Estimates With Time-Varying Exposures. *Am J Epidemiol* 2019;**188**: 231-8.

## Figure Legend

Figure 1. Illustration of the total effect and direct effect of BMI on lung cancer. The direct effect of BMI on lung cancer is the effect BMI has on lung cancer not via any other exposure variables, which is equal to  $\beta_{XZ}$ . The total effect of BMI on lung cancer is the effect of BMI on lung cancer directly plus the effect of BMI on lung cancer via smoking, which is equal to  $\beta_{XZ} + \alpha_{XY}\beta_{YZ}$ .

Figure 2. Odds ratios (ORs) and 95% confidence intervals (CIs) for the effect of BMI and smoking phenotypes on lung cancer and histological subtypes estimated using multivariable Mendelian randomization (MR) inverse-variance weighted (IVW) approach. Overall: overall lung cancer; Adeno: lung adenocarcinoma; Squam: lung squamous cell carcinoma; Small: small cell lung cancer. Colors indicate five exposures. BMI, body mass index; SmkInit, smoking status; AgeSmk, age of smoking initiation; SmkCes, smoking cessation; CigDay, cigarettes per day. For example, there were 776 BMI-associated SNPs been used in multivariable MR jointly with other smoking-associated SNPs to investigate their causal effects on overall lung cancer. No significant effect of BMI on overall lung cancer was observed (OR = 0.95, 95% CI = 0.87-1.03,  $P = 0.224$ ).