



A multiethnic bidirectional Mendelian randomization study negates causal effects of C-reactive protein concentrations on lung cancer

Xiang-Rong Wu^{1,2#}, Hao-Xin Peng^{1,2#}, Shan Xiong^{1#}, Yao-Kai Wen^{3,4#}, Jia-Na Chen^{1,2}, Cai-Chen Li¹, Yu Jiang^{1,2}, Zi-Xuan Su^{1,2}, Jun Liu¹, Jian-Xing He¹, Wen-Hua Liang^{1,5}, Xiu-Yu Cai⁶

¹Department of Thoracic Surgery and Oncology, the First Affiliated Hospital of Guangzhou Medical University, State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, Guangzhou, China; ²Nanshan School, Guangzhou Medical University, Guangzhou, China; ³School of Medicine, Tongji University, Shanghai, China; ⁴Department of Medical Oncology, Shanghai Pulmonary Hospital, Tongji University Medical School Cancer Institute, Tongji University School of Medicine, Shanghai, China; ⁵Department of Medical Oncology, The First People's Hospital of Zhaoqing, Zhaoqing, China; ⁶Department of General Internal Medicine, Sun Yat-sen University Cancer Centre, State Key Laboratory of Oncology in South China, Collaborative Innovation Centre for Cancer Medicine, Guangzhou, China

Contributions: (I) Conception and design: XR Wu, HX Peng, S Xiong, YK Wen; (II) Administrative support: J Liu, JX He, WH Liang, XY Cai; (III) Provision of study materials or patients: JN Chen, CC Li, Y Jiang, ZX Su; (IV) Collection and assembly of data: JN Chen, CC Li, Y Jiang, ZX Su; (V) Data analysis and interpretation: XR Wu, HX Peng, S Xiong, YK Wen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Jian-Xing He, MD. Department of Thoracic Surgery and Oncology, The First Affiliated Hospital of Guangzhou Medical University, State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, Guangzhou 510120, China. Email: drjianxing.he@gmail.com; Wen-Hua Liang, MD. Department of Thoracic Surgery and Oncology, The First Affiliated Hospital of Guangzhou Medical University, State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, Guangzhou 510120, China; Department of Medical Oncology, The First People's Hospital of Zhaoqing, Zhaoqing 526020, China. Email: liangwh1987@163.com; Xiu-Yu Cai. Department of General Internal Medicine, Sun Yat-sen University Cancer Centre, State Key Laboratory of Oncology in South China, Collaborative Innovation Centre for Cancer Medicine, Guangzhou 510060, China. Email: caixy@sysucc.org.cn.

Background: Significantly rising plasma circulating C-reaction protein (CRP) concentrations are pervasive in lung cancer (LC) development, demonstrating a bidirectional relation. However, it remains uncertain whether the causation between them exists, and the degree to which the effect varies across different ethnic ancestries remains unknown. Therefore, we attempted to investigate the causal relationship between these two phenotypes.

Methods: With summary statistics of CRP-related single nucleotide polymorphisms (SNPs) identified by several large-scale genome-wide association studies (GWAS) datasets based on five ethnic ancestries coverage worldwide, we implemented bidirectional two-sample Mendelian randomization (MR) analyses. Genetic summary data of 11,348 LC cases and 15,861 controls from the International Lung Cancer Consortium (ILCCO) were applied. The inverse-variance weighted (IVW) approach was utilized as the principal analysis, supplemented by various complementary methods.

Results: MR study did not reveal the causal relationship shared across genetically predisposed CRP blood concentrations and LC risk (OR =1.022, 95% CI: 0.965–1.083, P=0.455) including pathological subtypes (OR =1.026, 95% CI: 0.947–1.112, P=0.534 for lung adenocarcinoma; OR =1.060, 95% CI: 0.970–1.158, P=0.201 for squamous cell lung cancer). Further analyses among East Asian, Hispanic/Latin American, European, African American/Afro-Caribbean, and South Asian populations revealed consistent null causation. Additionally, the causal effects of LC on CRP concentrations were not statically significant (OR =0.999, 95% CI: 0.977–1.021, P=0.923).

Conclusions: We did not observe a bidirectional causal association between CRP blood concentrations on LC among East Asian, Hispanic/Latin American, European, African American/Afro-Caribbean, and South Asian ancestry individuals.

Keywords: Lung cancer (LC); C-reactive protein (CRP); Mendelian randomization; risk factor; genome-wide association study

Submitted Sep 13, 2021. Accepted for publication Nov 08, 2021.

doi: 10.21037/tlcr-21-750

View this article at: <https://dx.doi.org/10.21037/tlcr-21-750>

Introduction

Cancer is a leading global public health issue and the second major cause of death in the United States. Particularly, lung cancer (LC) is currently the second most common cancer and the major cause of cancer-related death globally (1). The American Cancer Society estimates that 235,760 new LC cases and 131,880 LC deaths are projected to occur in the United States in 2021 (2). Thus, early recognition of potentially modifiable risk factors contributes to better prevention of LC.

Growing studies have demonstrated that chronic inflammation, especially within the respiratory system, might be a risk factor for LC (3-6). Several inflammatory conditions, such as chronic pulmonary infection (7) and chronic obstructive pulmonary disease (8), were correlated with an increased risk of LC. It has been proposed that LC could be induced by pulmonary inflammation through several pathways, such as the increase in angiogenesis during the repair of the damaged tissue and the generation of reactive oxygen (9). Simultaneously, circulating C-reactive protein (CRP) was considered as a systemic marker of chronic inflammation (10), and elevated CRP blood concentrations have been widely evaluated for their association with LC (3,11-20). Up till now, two assumptions have been presented to explain the mechanisms between elevated CRP concentrations and the risk of LC (10). On the one hand, elevated CRP concentrations and chronic inflammation are causally related to the carcinogenesis of the lung indeed. On the other hand, CRP concentrations may also indicate a systemic response to ongoing disease or comorbidity, such as an occult cancer or a premalignant state. The majority of previous cohort studies and a meta-analysis (21) denoted that elevated CRP concentrations were correlated with LC occurrence. However, given that cigarette smoking is by far the leading risk factor for LC, CRP is widely seen

as a predictor of LC in smokers and former smokers (22), while cigarette smoking itself has been reported to increase circulating levels of CRP directly (23). Hence, conventional observational studies are prone to bias by potential confounding factors or reverse causation, and the causality was underpowered for definitive conclusions from previous studies.

Using genetic variations as instrumental variables (IVs), Mendelian randomization (MR) analysis is a novel epidemiology method to estimate the causation between an exposure and an outcome, with less impressionability to reverse causation and unmeasured confounders. In settings for which the IV assumptions are well justified, the findings could help optimize drug development or clinical trials and inform clinical or public health decision-making (24). Single nucleotide polymorphism (SNP) is a single base-pair difference in the DNA sequence of individuals within a species, which is the most common type of genetic variation in humans (25). Given that genetic variants are allocated at conception randomly, they are generally independent of environmental risk factors and precede the risk factors and the diseases' onset (26). As for CRP, it has been estimated that the heritability of CRP blood concentrations was estimated from 25% to 40% (27), suggesting that genetic architecture might contribute to modulating CRP concentrations. Meanwhile, based on the summary data from genome-wide association studies (GWAS), our two-sample MR analysis could assess their causal relationships more comprehensively with robust statistical power (28,29). Previous MR studies (5,30-32) have illustrated that no causation was detected between genetically regulated elevated CRP concentrations and risk of LC. However, the sample size of LC patients in previous MR studies was relatively limited, insufficient to provide adequate statistical power to evaluate their causal nexus. Then, due to the limitation of the GWASs at the time, causal inferences

provided by previous studies solely applied to populations of European ancestry; therefore, the findings were lack of generalizable value. Moreover, since genetically higher LC risk can influence CRP concentrations, no bidirectional MR study has been conducted, and therefore, previous findings were likely to be biased by reverse causality.

Generally, using the recent largescale meta-analysis of the GWASs updating CRP-specific SNPs among diverse populations, the present study could offer the latest comprehensive evidence for assessing the bidirectional causal effect between genetically regulated CRP concentrations and risk of LC with the two-sample MR method.

Methods

Study participants of lung cancer

In general, 11,348 LC cases and 15,861 controls from the International Lung Cancer Consortium (ILCCO) (33), an international research group that established and implemented LC research from different geographical areas and ethnicities, were used as epidemiological summary-level data. Subgroup analyses were stratified according to the pathology classification of non-small cell lung cancer (NSCLC). Specifically, we divided types of cancer into lung adenocarcinoma (LUAD) (3442 cases, 14,894 controls) as well as squamous cell lung cancer (LUSC) (3275 cases, 15,038 controls). The study was conducted following the Declaration of Helsinki (as revised in 2013) (34).

Genetic variants associated with elevated CRP concentrations

The latest large-scale GWAS datasets were retrieved from MR-Base, an open GWAS project developed at the MRC Integrative Epidemiology Unit (IEU) at the University of Bristol (35). Through Bonferroni correction to control the family-wise error rate (FWER), the $P=5\times 10^{-8}$ threshold was widely accepted for association identification between a common genetic variant and a trait of interest, given the linkage disequilibrium (LD) structure of the genome (36). Using the MR-Base platform, SNPs associated with elevated CRP concentrations were initially selected from Neale Lab, Pan-UK Biobank (UKB) team, RIKEN Center for Integrative Medical Sciences, and the European Bioinformatics Institute (EBI) database of complete GWAS summary data at the genome-wide significance

threshold ($P<5\times 10^{-8}$). Utilizing LD analysis, we attempted to exclude SNPs once mutual LD surpassed the limited value ($R^2<0.001$) with a larger P value conjugately. Eventually, the final IV set was established, including 204 SNPs (available online: <https://cdn.amegroups.com/static/public/tlcr-21-750-01.xlsx>). Among them, 174 SNPs were of European ancestry, 13 SNPs were of Hispanic or Latin American origin, 7 SNPs were of South Asian ancestry, 6 SNPs were of East Asian ancestry, and 4 SNPs were of African American or Afro-Caribbean ancestry. Risk alleles and baseline alleles were encoded separately in accordance with the association with the rate of CRP concentrations. As the sample size of 204 instruments and 310,305 individual samples in our study, the F-statistic was 3,135.39 as estimated (37) given the type I error rate =0.05, which suggested noticeable correlativity for the MR analysis.

Statistical analysis

MR method was implied as the statistical analysis tool, which possesses three suppositions as foundations (38): (I) the IVs are robustly correlating with elevated CRP concentrations; (II) the IVs affect LC merely via their effect on elevated CRP concentrations (i.e., the IVs are independent of the outcome given the exposure), and (III) the IVs are independent of any confounders. As demonstrated by the previous studies (35) and the screening procedure of SNPs described in the previous paragraph, the first assumption was met.

Two-sample MR analyses were conducted to evaluate the potential bidirectional associations between elevated CRP concentrations and LC risk (Figure 1). In addition to the aggregate effect, causal inferences were obtained from five independent ancestries for specific evaluation among different ethnic groups. Five well-established MR analytical methods were applied, including inverse-variance weighted (IVW), weighted median, MR-Egger, weighted mode, and simple mode. IVW was applied to the combination of multiple IVs as sole estimation of genetic variants by weighted score, which has the most statistical power among the five methods (39). The rest of the methods were performed for sensitivity analysis and to indirectly test the second assumption. Global pleiotropic effects were obtained from the MR-Egger analyses based on the intercept. Cochran's Q statistic and I^2 were inspected to estimate heterogeneity. Furthermore, we conducted leave-one-out analyses to investigate whether the estimation of MR was determined or biased by an individual SNP

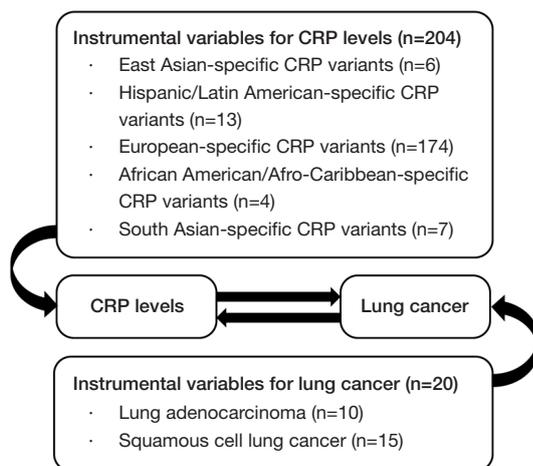


Figure 1 The overview of our study design aiming at revealing the bidirectional relations between C-reactive protein (CRP) and lung cancer. Two hundred and four CRP-specific single nucleotide polymorphisms (SNPs) (originated from East Asian, Hispanic/Latin American, European, African American/Afro-Caribbean, and South Asian ancestries), and 20 lung cancer-specific SNPs (including lung adenocarcinoma and squamous cell lung cancer) were applied as instrumental variables respectively to evaluate the bidirectional causal relation.

by omitting a single SNP successively. MR analyses were performed in R (version 3.6.2) using the package TwoSampleMR (version 0.5.0) (35).

Results

The basic characteristics and their *F* statistics for selected summary level GWASs applied in our MR study are given in *Table 1*, and the detailed information of each SNP used in our study is listed in table available online: <https://cdn.amegroups.cn/static/public/tlcr-21-750-01.xlsx>.

In the MR analysis of mixed ancestries using 204 SNPs, the outcome of IVW method demonstrated that although genetically regulated one-unit increased in the log-transformed CRP concentrations were associated with a relatively 2.2% increased risk of LC, no statistically meaningful conclusions could be drawn (OR =1.022, 95% CI: 0.965–1.083, *P*=0.455) (*Figure 2*), which was also consistent regarding pathological subtypes including LUAD (OR =1.026, 95% CI: 0.947–1.112, *P*=0.534) and LUSC (OR =1.060, 95% CI: 0.970–1.158, *P*=0.201). Findings were consistent among sensitivity analyses, demonstrating null causal relationship between both phenotypes (available

online: <https://cdn.amegroups.cn/static/public/tlcr-21-750-02.xlsx>). Meanwhile, as for the causal effect among different ethnic populations, no association reached statistical significance among East Asian (OR =1.075, 95% CI: 0.754–1.532, *P*=0.690), Hispanic/Latin American (OR =0.998, 95% CI: 0.884–1.127, *P*=0.980), European (OR =1.044, 95% CI: 0.945–1.155, *P*=0.397), African American/Afro-Caribbean (OR =1.055, 95% CI: 0.884–1.260, *P*=0.551), and South Asian population (OR =0.999, 95% CI: 0.911–1.096, *P*=0.984). Detailed results on LC subtypes of ethnic-specific analyses are presented in *Figure 3*.

Regarding the impact of LC upon CRP concentrations, similar modest and nonsignificant estimates of causal effect were observed (OR =0.999, 95% CI: 0.977–1.021, *P*=0.923), conforming the absence of bidirectional effects. Subtypes including LUSC (OR =0.996, 95% CI: 0.970–1.023, *P*=0.790) and LUAD (OR =1.011, 95% CI: 0.991–1.031, *P*=0.300) supported the same conclusion. Lack of causal association also predominated among population subgroups (*Figure 3*).

The results of sensitivity analyses among all study populations are shown in table available online: <https://cdn.amegroups.cn/static/public/tlcr-21-750-02.xlsx>. Using MR-Egger regression, our findings did not support the existence of global pleiotropic assumptions in most of the study outcomes (*Table S1*), while evidence of heterogeneity was found in LC effect on European-specific CRP variants (*Table S2*). Leave-one-out studies of the LC overall and subgroup analyses showed no evidence that a single SNP had an impact upon the overall effect of CRP variants on LC risks (available online: <https://cdn.amegroups.cn/static/public/tlcr-21-750-03.xlsx>).

Discussion

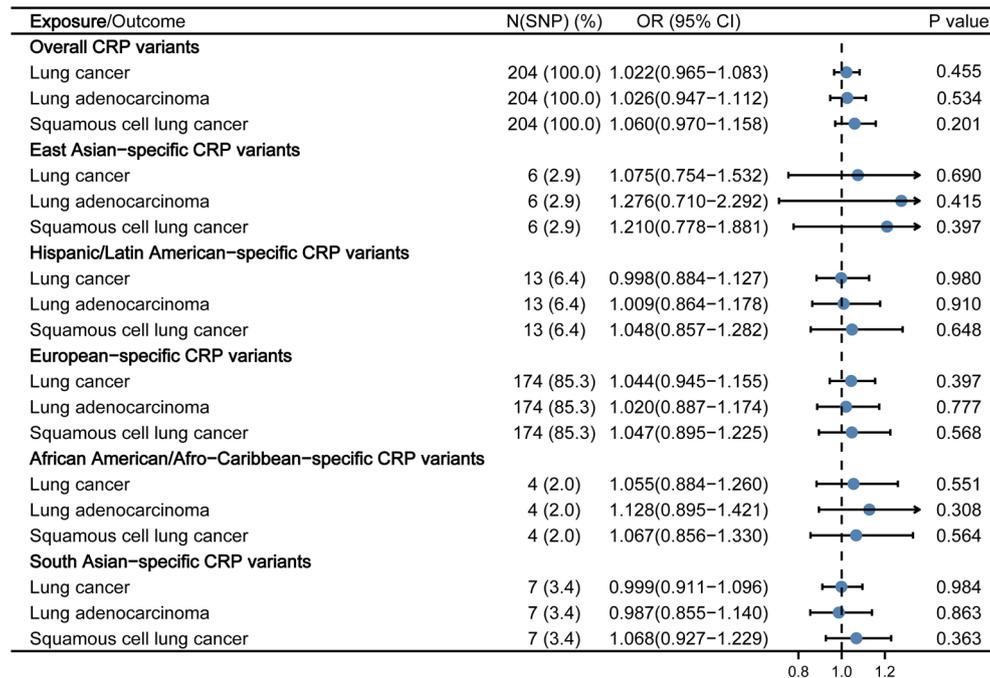
This is the first study to conduct bidirectional MR analyses between CRP concentrations and LC among different ethnic backgrounds. No bidirectional causation between genetically regulated elevated concentrations of CRP and LC risk was observed in our study among East Asian (*P*=0.690, nSNP =6, n=75,391), Hispanic/Latin American (*P*=0.980, nSNP =13, n=15,912), European (*P*=0.397, nSNP =174, n=204,402), African American/Afro-Caribbean (*P*=0.551, nSNP =4, n=6,203), and South Asian populations (*P*=0.984, nSNP =7, n=8,397). Additionally, no causality was found in subgroup analyses concerning pathological types as well.

The results are inconsistent with most previous prospective cohort studies (3,11,12,14,16-19) using

Table 1 The basic characteristics and their individual F statistics for selected summary level GWASs applied in our MR study.

Trait	Sample size	Sex	First author/consortium	F statistics
Overall C-reactive protein (CRP)	310,305	Males and females	–	3,135.39
East Asian-specific CRP variants	75,391	Males and females	Ishigaki K	762.53
Hispanic/Latin American-specific CRP variants	15,912	Males and females	Wojcik GL	161.73
European-specific CRP variants	204,402	Males and females	Neale lab	2,065.67
African American/Afro-Caribbean-specific CRP variants	6,203	Males and females	Pan-UKB team	63.66
South Asian-specific CRP variants	8,397	Males and females	Pan-UKB team	85.82
Lung cancer	27,209	Males and females	ILCCO	275.84
Lung adenocarcinoma	18,336	Males and females	ILCCO	186.21
Squamous cell lung cancer	18,313	Males and females	ILCCO	185.98

ILCCO, the International Lung Cancer Consortium.

**Figure 2** Inverse-variance weighted results of MR study investigating the causal association between genetically regulated CRP levels on lung cancer among different ethnic populations. MR, Mendelian randomization; CRP, C-reactive protein.

measured plasma concentrations of circulating CRP to explore the risk of LC. Shiels *et al.* (16) demonstrated that elevated measured CRP concentrations were related to the risk of LC in two independent studies from the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. Allin *et al.* (11) also illustrated that elevated plasma concentrations of CRP in cancer-free individuals were

relevant to an increased risk of LC. Besides, the latest meta-analysis (21) denoted that elevated plasma concentrations of CRP were also related to an increased risk of LC. Interestingly, after a stratified analysis (15) based on smoking status, the positive association between circulating high sensitivity CRP concentrations and the risk of LC disappeared when considering non-smokers. Thus, rather

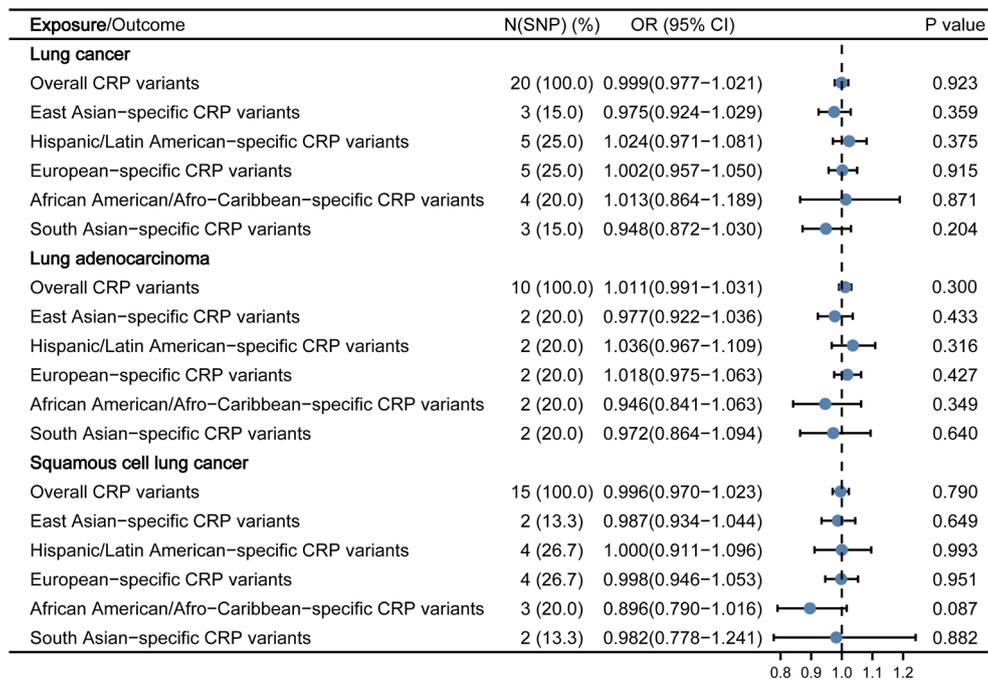


Figure 3 Inverse-variance weighted results of bidirectional MR study investigating the causal association between lung cancer and CRP levels among different ethnic populations. MR, Mendelian randomization; CRP, C-reactive protein.

than a causal factor, the authors (15) concluded that circulating high sensitivity CRP concentration could be a pre-diagnostic marker of LC. Since traditional observational studies are vulnerable to potential confounding factors or reverse causation, we supposed that previous results might be influenced by underlying confounders (e.g., cigarette smoking status and older age stratifications).

Our study showed similar results in effect directions compared with previous MR studies (5,30-32), whereas these studies came to the conclusions from a limited LC sample size. Among them, the largest sample size of LC patients was only 416 (30). Therefore, they might not be precise enough to evaluate causation between genetically elevated CRP concentrations and LC risk due to the deficiency of statistical power. In comparison, we specifically selected a total of 27,209 participants from ILCCO with 11,348 LC cases and 15,861 controls as our study participants, which were much more extensive than previous ones and offered adequate statistical power. Meanwhile, two (5,30) of them didn't identify SNPs through GWAS, and the majority of chosen SNPs haven't been confirmed by subsequent GWAS (27,40,41), indicating those SNPs might not be authentically associated with elevated CRP concentrations. Even for the other two

studies (31,32) with GWAS-identified SNPs, some of their chosen SNPs were also proposed to have nothing to do with elevated CRP concentrations subsequently (27,40,41), such as rs4903031, rs6901250, and rs4705952. Additionally, genetic pleiotropy could occur without sensitivity analyses, indicating previously chosen SNPs might relate to other inflammatory processes and result in a spurious association. Consequently, biased estimation for the causation could happen due to those irrelevant SNPs and potential global pleiotropic effects in previous MR studies. Furthermore, more relevant SNPs have been updated in recent years. Our study included 204 SNPs from the latest GWAS as the IV set, which would explain the variance of elevated CRP concentrations more comprehensively and precisely. Generally, our updated MR study provided the latest and much more substantial evidence for evaluating the causality between genetically regulated elevated CRP concentrations and LC risk.

Regarding the etiology of LC, increasing evidence have illustrated that chronic inflammation may play a significant role amid the carcinogenesis process (4-7,16,42,43). Inflammatory conditions, such as chronic obstructive pulmonary disease (3,44-47), chlamydia pneumonia (4,48), tuberculosis (6), and human immunodeficiency virus (HIV) (49) seem to correlate

with an increased LC risk. Through the nuclear factor kappa B (NF- κ B) pathway, chronic inflammation may actively promote the malignant process, the increase in angiogenesis and proliferation of cells during tissue repair, and the up-regulation of antiapoptotic genes through reactive oxygen and nitrogen species production (9). Considered a systemic marker of chronic inflammation (5), elevated plasma CRP concentrations were associated with an increased LC risk in many conventional observational studies. However, our MR analysis denoted that elevated CRP concentrations were not causally associated with LC risk, which could be explained from two aspects. For one thing, chronic inflammation promotes the carcinogenesis of the lung, possibly by means of other inflammatory mediators except for CRP. It has been reported that soluble tumor necrosis factor receptor-2 (sTNFR2), serum amyloid A (SAA), and monokine induced by gamma interferon (CXCL9/MIG) were also related to LC risk (16). For another, due to the natural defects of observational studies, their results could be influenced by unmeasured confounders, such as smoking status and elder age. In addition, inflammation and elevated CRP concentrations could be induced by tumor growth, and CRP could also be an indicator of an immune response to tumor antigens (10). Consequently, elevated CRP concentrations might also respond to occult or very early-stage cancer (16). However, CRP may not always act as a cancer promoter but sometimes play a role in treating cancer. Sasaki *et al.* observed that CRP can inhibit lymph node metastasis and lymphatic angiogenesis of squamous cell carcinoma through injecting CRP into mice model (50). As for the prognostic effects of CRP, Shinohara and colleagues found high serum circulating CRP levels on postoperative days were associated with enhanced 5-year overall survival (OS) as well as recurrence-free survival (RFS) in NSCLC patients (51). Instead, Okada *et al.* reported the high serum CRP levels during perioperative period were a poor prognostic factor for OS and RFS in NSCLC patients (52). Consequently, the impact of CRP levels on the incidence and prognosis of LC was nuanced and complex, and further studies with larger sample sizes and better designs are in an urgent need.

Several merits deserve explicit mention. First, this is the first study to assess the bidirectional causation between elevated CRP concentrations and LC risk to date. Second, we used the most comprehensive GWAS-identified SNPs as an IV set. Compared with the previous studies, our study with 204 SNPs explained more variance of elevated CRP concentrations. Thanks to the robustly associated IVs (F -statistics = 3,135.39) and the great sample size ($n=310,305$), our study was capable of providing a relatively

accurate evaluation of a causal inference. Third, to the best of our knowledge, this is the first MR study that conducted stratification analyses according to different ethnic populations. Thus, we were able to investigate whether their association could vary under different ancestry backgrounds.

Notwithstanding the advantages, several limitations in our study cannot be neglected. First, even though the most comprehensive set of multiethnic genetic variants were applied for our research so far, only a part of the variance of elevated CRP concentrations could be explained in the population. Some unknown CRP-related SNPs, which deserve further investigation in future GWASs, could play an essential role in LC development. Besides, due to the methodology limitations currently, MR assumptions could not be thoroughly tested, and therefore potential violations against the assumptions may occur. To overcome this difficulty, instead of directivity evaluated in the second assumption, we implemented additional sensitivity analyses, which showed the absence of pleiotropic effects in our study, suggesting no violation of the second MR assumption. Moreover, depending on GWAS summary statistics, the two-sample MR methods assume a linear relationship between the exposure and the outcome. Hence, it is possible that a few genetic variants were also related to confounding factors of CRP and LC.

Overall, our study negates bidirectional causal effects of CRP concentrations on LC among East Asian, Hispanic/Latin American, European, African American/Afro-Caribbean, and South Asian populations. Still, it opens a new concept for the current research orientation. Further investigation of the profound relationship between both phenotypes is required to be unveiled in pathologic and biochemistry aspects.

Acknowledgments

Funding: This work was supported by the National Key R&D Program of China (2016YFC0905400), China National Science Foundation (81871893, 81501996), Cultivation of Guangdong College Students' Scientific and Technological Innovation ("Climbing Program" Special Funds) (Grant number pdjh2020a0480 and pdjh2021a0407), Key Project of Guangzhou Scientific Research Project (201804020030), High-level university construction project of Guangzhou Medical University (20182737, 201721007, 201715907, 2017160107), IVATS National key R&D Program (2017YFC0907903, 2017YFC0112704)

and Application, Industrialization and Generalization of Surgical Incision Protector (2011B090400589).

Footnote

Peer Review File: Available at <https://dx.doi.org/10.21037/tlcr-21-750>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/tlcr-21-750>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted following the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71:209-49.
- Siegel RL, Miller KD, Fuchs HE, et al. Cancer Statistics, 2021. *CA Cancer J Clin* 2021;71:7-33. Erratum in: *CA Cancer J Clin*. 2021;71:359.
- Allin KH, Bojesen SE, Johansen JS, et al. Cancer risk by combined levels of YKL-40 and C-reactive protein in the general population. *Br J Cancer* 2012;106:199-205.
- Chaturvedi AK, Gaydos CA, Agreda P, et al. Chlamydia pneumoniae infection and risk for lung cancer. *Cancer Epidemiol Biomarkers Prev* 2010;19:1498-505.
- Chaturvedi AK, Caporaso NE, Katki HA, et al. C-reactive protein and risk of lung cancer. *J Clin Oncol* 2010;28:2719-26.
- Engels EA, Wu X, Gu J, et al. Systematic evaluation of genetic variants in the inflammation pathway and risk of lung cancer. *Cancer Res* 2007;67:6520-7.
- Shiels MS, Albanes D, Virtamo J, et al. Increased risk of lung cancer in men with tuberculosis in the alpha-tocopherol, beta-carotene cancer prevention study. *Cancer Epidemiol Biomarkers Prev* 2011;20:672-8.
- Koshiol J, Rotunno M, Consonni D, et al. Chronic obstructive pulmonary disease and altered risk of lung cancer in a population-based case-control study. *PLoS One* 2009;4:e7380.
- Ballaz S, Mulshine JL. The potential contributions of chronic inflammation to lung carcinogenesis. *Clin Lung Cancer* 2003;5:46-62.
- Heikkilä K, Ebrahim S, Lawlor DA. A systematic review of the association between circulating concentrations of C reactive protein and cancer. *J Epidemiol Community Health* 2007;61:824-33.
- Allin KH, Bojesen SE, Nordestgaard BG. Baseline C-reactive protein is associated with incident cancer and survival in patients with cancer. *J Clin Oncol* 2009;27:2217-24.
- Allin KH, Nordestgaard BG. Elevated C-reactive protein in the diagnosis, prognosis, and cause of cancer. *Crit Rev Clin Lab Sci* 2011;48:155-70.
- dos Santos Silva I, De Stavola BL, Pizzi C, et al. Circulating levels of coagulation and inflammation markers and cancer risks: individual participant analysis of data from three long-term cohorts. *Int J Epidemiol* 2010;39:699-709.
- Il'yasova D, Colbert LH, Harris TB, et al. Circulating levels of inflammatory markers and cancer risk in the health aging and body composition cohort. *Cancer Epidemiol Biomarkers Prev* 2005;14:2413-8.
- Muller DC, Larose TL, Hodge A, et al. Circulating high sensitivity C reactive protein concentrations and risk of lung cancer: nested case-control study within Lung Cancer Cohort Consortium. *BMJ* 2019;364:k4981.
- Shiels MS, Katki HA, Hildesheim A, et al. Circulating Inflammation Markers, Risk of Lung Cancer, and Utility for Risk Stratification. *J Natl Cancer Inst* 2015;107:djv199.
- Shiels MS, Pfeiffer RM, Hildesheim A, et al. Circulating inflammation markers and prospective risk for lung cancer. *J Natl Cancer Inst* 2013;105:1871-80.
- Siemes C, Visser LE, Coebergh JW, et al. C-reactive protein levels, variation in the C-reactive protein gene, and cancer risk: the Rotterdam Study. *J Clin Oncol* 2006;24:5216-22.

19. Van Hemelrijck M, Holmberg L, Garmo H, et al. Association between levels of C-reactive protein and leukocytes and cancer: three repeated measurements in the Swedish AMORIS study. *Cancer Epidemiol Biomarkers Prev* 2011;20:428-37.
20. Heikkilä K, Harris R, Lowe G, et al. Associations of circulating C-reactive protein and interleukin-6 with cancer risk: findings from two prospective cohorts and a meta-analysis. *Cancer Causes Control* 2009;20:15-26.
21. Guo YZ, Pan L, Du CJ, et al. Association between C-reactive protein and risk of cancer: a meta-analysis of prospective cohort studies. *Asian Pac J Cancer Prev* 2013;14:243-8.
22. Tonstad S, Cowan JL. C-reactive protein as a predictor of disease in smokers and former smokers: a review. *Int J Clin Pract* 2009;63:1634-41.
23. Gallus S, Lugo A, Suatoni P, et al. Effect of Tobacco Smoking Cessation on C-Reactive Protein Levels in A Cohort of Low-Dose Computed Tomography Screening Participants. *Sci Rep* 2018;8:12908.
24. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018;362:k601.
25. Schork NJ, Fallin D, Lanchbury JS. Single nucleotide polymorphisms and the future of genetic epidemiology. *Clin Genet* 2000;58:250-64.
26. Lawlor DA, Harbord RM, Sterne JA, et al. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008;27:1133-63.
27. Dehghan A, Dupuis J, Barbalic M, et al. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation* 2011;123:731-8.
28. Burgess S, Scott RA, Timpson NJ, et al. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol* 2015;30:543-52.
29. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol* 2013;178:1177-84.
30. Allin KH, Nordestgaard BG, Zacho J, et al. C-reactive protein and the risk of cancer: a mendelian randomization study. *J Natl Cancer Inst* 2010;102:202-6.
31. Heikkilä K, Silander K, Salomaa V, et al. C-reactive protein-associated genetic variants and cancer risk: findings from FINRISK 1992, FINRISK 1997 and Health 2000 studies. *Eur J Cancer* 2011;47:404-12.
32. Prizment AE, Folsom AR, Dreyfus J, et al. Plasma C-reactive protein, genetic risk score, and risk of common cancers in the Atherosclerosis Risk in Communities study. *Cancer Causes Control* 2013;24:2077-87.
33. Wang Y, McKay JD, Rafnar T, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. *Nat Genet* 2014;46:736-41.
34. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013;310:2191-4.
35. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018;7:34408.
36. Chen Z, Boehnke M, Wen X, et al. Revisiting the genome-wide significance threshold for common variant GWAS. *G3 (Bethesda)* 2021;11:jkaa056.
37. Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011;40:755-64.
38. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, et al. Methodological challenges in mendelian randomization. *Epidemiology* 2014;25:427-35.
39. Noyce AJ, Kia DA, Hemani G, et al. Estimating the causal influence of body mass index on risk of Parkinson disease: A Mendelian randomisation study. *PLoS Med* 2017;14:e1002314.
40. Ligthart S, Vaez A, Vösa U, et al. Genome Analyses of >200,000 Individuals Identify 58 Loci for Chronic Inflammation and Highlight Pathways that Link Inflammation and Complex Disorders. *Am J Hum Genet* 2018;103:691-706.
41. Kocarnik JM, Pendergrass SA, Carty CL, et al. Multiancestral analysis of inflammation-related genetic variants and C-reactive protein in the population architecture using genomics and epidemiology study. *Circ Cardiovasc Genet* 2014;7:178-88.
42. Pine SR, Mechanic LE, Enewold L, et al. Increased levels of circulating interleukin 6, interleukin 8, C-reactive protein, and risk of lung cancer. *J Natl Cancer Inst* 2011;103:1112-22.
43. Shiels MS, Engels EA, Shi J, et al. Genetic variation in innate immunity and inflammation pathways associated with lung cancer risk. *Cancer* 2012;118:5630-6.
44. Santillan AA, Camargo CA Jr, Colditz GA. A meta-analysis of asthma and risk of lung cancer (United States). *Cancer Causes Control* 2003;14:327-34.

45. Sin DD, Man SF. Impact of cancers and cardiovascular diseases in chronic obstructive pulmonary disease. *Curr Opin Pulm Med* 2008;14:115-21.
46. Purdue MP, Gold L, Järholm B, et al. Impaired lung function and lung cancer incidence in a cohort of Swedish construction workers. *Thorax* 2007;62:51-6.
47. Brody JS, Spira A. State of the art. Chronic obstructive pulmonary disease, inflammation, and lung cancer. *Proc Am Thorac Soc* 2006;3:535-7.
48. Littman AJ, Jackson LA, Vaughan TL. Chlamydia pneumoniae and lung cancer: epidemiologic evidence. *Cancer Epidemiol Biomarkers Prev* 2005;14:773-8.
49. Engels EA, Brock MV, Chen J, et al. Elevated incidence of lung cancer among HIV-infected individuals. *J Clin Oncol* 2006;24:1383-8.
50. Sasaki T, Motoyama S, Sato Y, et al. C-reactive protein inhibits lymphangiogenesis and resultant lymph node metastasis of squamous cell carcinoma in mice. *Surgery* 2013;154:1087-92.
51. Shinohara S, Sugaya M, Onitsuka T, et al. Prognostic Impact of Postoperative C-reactive Protein for Non-small Cell Lung Cancer Following Lobectomy. *Anticancer Res* 2018;38:3193-8.
52. Okada S, Shimomura M, Tsunozuka H, et al. Prognostic Significance of Perioperative C-Reactive Protein in Resected Non-Small Cell Lung Cancer. *Semin Thorac Cardiovasc Surg* 2020;32:1046-55.

Cite this article as: Wu XR, Peng HX, Xiong S, Wen YK, Chen JN, Li CC, Jiang Y, Su ZX, Liu J, He JX, Liang WH, Cai XY. A multiethnic bidirectional Mendelian randomization study negates causal effects of C-reactive protein concentrations on lung cancer. *Transl Lung Cancer Res* 2021;10(12):4380-4389. doi: 10.21037/tlcr-21-750

Table S1 MR-Egger regression analyses on detecting directional pleiotropy of Mendelian randomization.

Outcome	Exposure	Egger intercept	SE	P value
Lung adenocarcinoma	C-reactive protein	0.0013	0.0031	0.6907
Lung cancer	C-reactive protein	0.0005	0.0023	0.8411
Squamous cell lung cancer	C-reactive protein	-0.0018	0.0035	0.6039
Lung adenocarcinoma	East Asian-specific CRP variants	0.0010	0.0509	0.9850
Lung cancer	East Asian-specific CRP variants	-0.0384	0.0247	0.1945
Squamous cell lung cancer	East Asian-specific CRP variants	-0.0054	0.0352	0.8860
Lung adenocarcinoma	Hispanic/Latin American-specific CRP variants	0.0107	0.0179	0.5618
Lung cancer	Hispanic/Latin American-specific CRP variants	-0.0013	0.0144	0.9278
Squamous cell lung cancer	Hispanic/Latin American-specific CRP variants	-0.0037	0.0237	0.8774
Lung adenocarcinoma	European-specific CRP variants	0.0017	0.0037	0.6500
Lung cancer	European-specific CRP variants	0.0010	0.0027	0.7013
Squamous cell lung cancer	European-specific CRP variants	-0.0020	0.0042	0.6317
Lung adenocarcinoma	African American/Afro-Caribbean-specific CRP variants	-0.1058	0.0686	0.2632
Lung cancer	African American/Afro-Caribbean-specific CRP variants	-0.0878	0.0448	0.1891
Squamous cell lung cancer	African American/Afro-Caribbean-specific CRP variants	0.0553	0.0650	0.4844
Lung adenocarcinoma	South Asian-specific CRP variants	-0.0107	0.0238	0.6722
Lung cancer	South Asian-specific CRP variants	-0.0236	0.0152	0.1829
Squamous cell lung cancer	South Asian-specific CRP variants	-0.0068	0.0234	0.7842
C-reactive protein levels	Lung adenocarcinoma	-0.0347	0.0120	0.0199
C-reactive protein levels	Lung cancer	-0.0274	0.0086	0.0052
C-reactive protein levels	Squamous cell lung cancer	-0.0113	0.0099	0.2749
East Asian-specific CRP variants	Lung adenocarcinoma	-	-	-
Hispanic/Latin American-specific CRP variants	Lung adenocarcinoma	-	-	-
European-specific CRP variants	Lung adenocarcinoma	-	-	-
African American/Afro-Caribbean-specific CRP variants	Lung adenocarcinoma	-	-	-
South Asian-specific CRP variants	Lung adenocarcinoma	-	-	-
East Asian-specific CRP variants	Lung cancer	-0.0605	0.0221	0.2233
Hispanic/Latin American-specific CRP variants	Lung cancer	-0.0380	0.0273	0.2589
European-specific CRP variants	Lung cancer	-0.0256	0.0181	0.2511
African American/Afro-Caribbean-specific CRP variants	Lung cancer	-0.1581	0.0993	0.2523
South Asian-specific CRP variants	Lung cancer	0.0425	0.0578	0.5963
East Asian-specific CRP variants	Squamous cell lung cancer	-	-	-
Hispanic/Latin American-specific CRP variants	Squamous cell lung cancer	-0.0389	0.0335	0.3654
European-specific CRP variants	Squamous cell lung cancer	-0.0094	0.0231	0.7238
African American/Afro-Caribbean-specific CRP variants	Squamous cell lung cancer	-0.0501	0.0590	0.5520
South Asian-specific CRP variants	Squamous cell lung cancer	-	-	-

Table S2 Heterogeneity test of bidirectional Mendelian randomization

Outcome	Exposure	Method	Q	Q df*	Q P value	I ²
Lung adenocarcinoma	Overall CRP variants	MR Egger	226.679	202	1.12E-01	10.89%
Lung adenocarcinoma	Overall CRP variants	Inverse variance weighted	226.857	203	1.20E-01	10.52%
Lung cancer	Overall CRP variants	MR Egger	282.244	202	1.64E-04	28.43%
Lung cancer	Overall CRP variants	Inverse variance weighted	282.300	203	1.94E-04	28.09%
Squamous cell lung cancer	Overall CRP variants	MR Egger	285.919	202	9.26E-05	29.35%
Squamous cell lung cancer	Overall CRP variants	Inverse variance weighted	286.301	203	1.05E-04	29.10%
Lung adenocarcinoma	East Asian-specific CRP variants	MR Egger	8.466	4	7.59E-02	52.75%
Lung adenocarcinoma	East Asian-specific CRP variants	Inverse variance weighted	8.467	5	1.32E-01	40.95%
Lung cancer	East Asian-specific CRP variants	MR Egger	4.539	4	3.38E-01	11.88%
Lung cancer	East Asian-specific CRP variants	Inverse variance weighted	7.289	5	2.00E-01	31.41%
Squamous cell lung cancer	East Asian-specific CRP variants	MR Egger	1.364	4	8.50E-01	0.00%
Squamous cell lung cancer	East Asian-specific CRP variants	Inverse variance weighted	1.388	5	9.26E-01	0.00%
Lung adenocarcinoma	Hispanic/Latin American-specific CRP variants	MR Egger	7.854	11	7.26E-01	0.00%
Lung adenocarcinoma	Hispanic/Latin American-specific CRP variants	Inverse variance weighted	8.212	12	7.68E-01	0.00%
Lung cancer	Hispanic/Latin American-specific CRP variants	MR Egger	17.355	11	9.78E-02	36.62%
Lung cancer	Hispanic/Latin American-specific CRP variants	Inverse variance weighted	17.369	12	1.36E-01	30.91%
Squamous cell lung cancer	Hispanic/Latin American-specific CRP variants	MR Egger	20.534	11	3.85E-02	46.43%
Squamous cell lung cancer	Hispanic/Latin American-specific CRP variants	Inverse variance weighted	20.580	12	5.69E-02	41.69%
Lung adenocarcinoma	European-specific CRP variants	MR Egger	202.690	172	5.47E-02	15.14%
Lung adenocarcinoma	European-specific CRP variants	Inverse variance weighted	202.934	173	5.94E-02	14.75%
Lung cancer	European-specific CRP variants	MR Egger	247.330	172	1.47E-04	30.46%
Lung cancer	European-specific CRP variants	Inverse variance weighted	247.542	173	1.73E-04	30.11%
Squamous cell lung cancer	European-specific CRP variants	MR Egger	259.036	172	1.98E-05	33.60%
Squamous cell lung cancer	European-specific CRP variants	Inverse variance weighted	259.383	173	2.31E-05	33.30%
Lung adenocarcinoma	African American/Afro-Caribbean-specific CRP variants	MR Egger	0.440	2	8.03E-01	0.00%
Lung adenocarcinoma	African American/Afro-Caribbean-specific CRP variants	Inverse variance weighted	2.816	3	4.21E-01	0.00%
Lung cancer	African American/Afro-Caribbean-specific CRP variants	MR Egger	0.396	2	8.20E-01	0.00%
Lung cancer	African American/Afro-Caribbean-specific CRP variants	Inverse variance weighted	4.237	3	2.37E-01	29.19%
Squamous cell lung cancer	African American/Afro-Caribbean-specific CRP variants	MR Egger	0.765	2	6.82E-01	0.00%
Squamous cell lung cancer	African American/Afro-Caribbean-specific CRP variants	Inverse variance weighted	1.489	3	6.85E-01	0.00%
Lung adenocarcinoma	South Asian-specific CRP variants	MR Egger	2.362	5	7.97E-01	0.00%
Lung adenocarcinoma	South Asian-specific CRP variants	Inverse variance weighted	2.564	6	8.61E-01	0.00%
Lung cancer	South Asian-specific CRP variants	MR Egger	2.493	5	7.78E-01	0.00%
Lung cancer	South Asian-specific CRP variants	Inverse variance weighted	4.881	6	5.59E-01	0.00%
Squamous cell lung cancer	South Asian-specific CRP variants	MR Egger	2.961	5	7.06E-01	0.00%
Squamous cell lung cancer	South Asian-specific CRP variants	Inverse variance weighted	3.044	6	8.03E-01	0.00%
Overall CRP variants levels	Lung adenocarcinoma	MR Egger	14.144	8	7.81E-02	43.44%
Overall CRP variants levels	Lung adenocarcinoma	Inverse variance weighted	29.003	9	6.47E-04	68.97%
Overall CRP variants levels	Lung cancer	MR Egger	59.026	18	2.94E-06	69.51%
Overall CRP variants levels	Lung cancer	Inverse variance weighted	92.240	19	1.33E-11	79.40%
Overall CRP variants levels	Squamous cell lung cancer	MR Egger	82.661	13	3.48E-12	84.27%
Overall CRP variants levels	Squamous cell lung cancer	Inverse variance weighted	90.924	14	2.54E-13	84.60%
East Asian-specific CRP variants	Lung adenocarcinoma	Inverse variance weighted	4.124	1	4.23E-02	75.75%
Hispanic/Latin American-specific CRP variants	Lung adenocarcinoma	Inverse variance weighted	2.133	1	1.44E-01	53.12%
European-specific CRP variants	Lung adenocarcinoma	Inverse variance weighted	11.145	1	8.43E-04	91.03%
African American/Afro-Caribbean-specific CRP variants	Lung adenocarcinoma	Inverse variance weighted	0.086	1	7.69E-01	0.00%
South Asian-specific CRP variants	Lung adenocarcinoma	Inverse variance weighted	1.955	1	1.62E-01	48.84%
East Asian-specific CRP variants	Lung cancer	MR Egger	0.679	1	4.10E-01	0.00%
East Asian-specific CRP variants	Lung cancer	Inverse variance weighted	8.149	2	1.70E-02	75.46%
Hispanic/Latin American-specific CRP variants	Lung cancer	MR Egger	3.413	3	3.32E-01	12.11%
Hispanic/Latin American-specific CRP variants	Lung cancer	Inverse variance weighted	5.609	4	2.30E-01	28.69%
European-specific CRP variants	Lung cancer	MR Egger	39.422	3	1.41E-08	92.39%
European-specific CRP variants	Lung cancer	Inverse variance weighted	65.854	4	1.70E-13	93.93%
African American/Afro-Caribbean-specific CRP variants	Lung cancer	MR Egger	2.419	2	2.98E-01	17.31%
African American/Afro-Caribbean-specific CRP variants	Lung cancer	Inverse variance weighted	5.484	3	1.40E-01	45.30%
South Asian-specific CRP variants	Lung cancer	MR Egger	0.272	1	6.02E-01	0.00%
South Asian-specific CRP variants	Lung cancer	Inverse variance weighted	0.813	2	6.66E-01	0.00%
East Asian-specific CRP variants	Squamous cell lung cancer	Inverse variance weighted	2.708	1	9.99E-02	63.07%
Hispanic/Latin American-specific CRP variants	Squamous cell lung cancer	MR Egger	6.508	2	3.86E-02	69.27%
Hispanic/Latin American-specific CRP variants	Squamous cell lung cancer	Inverse variance weighted	10.896	3	1.23E-02	72.47%
European-specific CRP variants	Squamous cell lung cancer	MR Egger	61.891	2	3.64E-14	96.77%
European-specific CRP variants	Squamous cell lung cancer	Inverse variance weighted	67.002	3	1.87E-14	95.52%
African American/Afro-Caribbean-specific CRP variants	Squamous cell lung cancer	MR Egger	0.709	1	4.00E-01	0.00%
African American/Afro-Caribbean-specific CRP variants	Squamous cell lung cancer	Inverse variance weighted	1.430	2	4.89E-01	0.00%
South Asian-specific CRP variants	Squamous cell lung cancer	Inverse variance weighted	5.638	1	1.76E-02	82.26%

df*, degree of freedom.