Serum 25-hydroxyvitamin D levels and risk of lung cancer and histologic types: a Mendelian randomization analysis of the HUNT study

Yi-Qian Sun¹, *, Ben M Brumpton², ³, ⁴, Carolina Bonilla⁴, Sarah J. Lewis⁴, Stephen Burgess⁵, ⁶, Frank Skorpen¹, Yue Chen⁷, Tom IL Nilsen⁸, ⁹, Pål Richard Romundstad⁸ and Xiao-Mei Mai⁸

¹Department of Clinical and Molecular Medicine (IKOM), NTNU, Norwegian University of Science and Technology, Trondheim, Norway.

²K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway.

³Department of Thoracic Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway.

⁴MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK.

⁵MRC Biostatistics Unit, University of Cambridge, Cambridge, UK.

⁶Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

⁷School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Canada.

⁸Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway.

⁹Clinic of Anesthesia and Intensive Care, St Olavs Hospital, Trondheim University Hospital, Trondheim, Norway.
“This is an author-submitted, peer-reviewed version of a manuscript that has been accepted for publication in the European Respiratory Journal, prior to copy-editing, formatting and typesetting. This version of the manuscript may not be duplicated or reproduced without prior permission from the copyright owner, the European Respiratory Society. The publisher is not responsible or liable for any errors or omissions in this version of the manuscript or in any version derived from it by any other parties. The final, copy-edited, published article, which is the version of record, is available without a subscription 18 months after the date of issue publication.”

*Corresponding author*

Yi-Qian Sun, PhD

Department of Clinical and Molecular Medicine (IKOM), NTNU, Norwegian University of Science and Technology, Trondheim, Norway

Tel: +47 72576121

Email: yi-qian.sun@ntnu.no

**Word count**

Abstract: 196

Text: 2772

**Keywords**

histologic types, lung cancer, Mendelian randomization, serum 25-hydroxyvitamin D [25(OH)D], single-nucleotide polymorphisms, vitamin D
For Twitter (≤120 characters):

Mendelian randomization study did not suggest causal association between serum 25-hydroxyvitamin D and lung cancer risk
Abstract

We aimed to investigate potential causal associations between serum 25-hydroxyvitamin D [25(OH)D] levels and incidence of lung cancer overall and histologic types.

We performed a Mendelian randomization (MR) analysis using a prospective cohort study in Norway, including 54580 individuals and 676 incident lung cancer cases. A 25(OH)D allele score was generated based on vitamin D-increasing alleles of rs2282679, rs12785878 and rs10741657. Hazard ratios (HRs) with 95% confidence intervals (CIs) for incidence of lung cancer and histologic types were estimated in relation to the allele score. Inverse-variance weighted method using summarized data of individual single-nucleotide polymorphisms was applied to calculate the MR estimates.

The allele score accounted for 3.4% of the variation in serum 25(OH)D levels. There was no association between the allele score and lung cancer incidence overall, with HR being 0.99 (95% CI 0.93 to 1.06) per allele score. A 25 nmol/L increase in genetically determined 25(OH)D level was not associated with the incidence of lung cancer overall (MR estimate HR 0.96, 95% CI 0.54 to 1.69) or any histologic type.

MR analysis did not suggest causal association between 25(OH)D levels and risk of lung cancer overall or histologic types in this population-based cohort study.
Introduction

Vitamin D has been suggested to have a number of anti-carcinogenic potentials, such as stimulating differentiation, inducing apoptosis, and inhibiting invasion and metastasis [1, 2]. Epidemiological studies of the associations between circulating vitamin D and various cancers have shown inconsistent results [3, 4]. Lung cancer has been the most common cancer type for several decades worldwide, and it is also the most deadly cancer [5]. The main histologic types of lung cancer are small cell lung cancer (SCLC), adenocarcinoma and squamous cell carcinoma [6]. Adenocarcinoma is the most common histologic type of lung cancer in many countries [7]. Unlike SCLC and squamous cell carcinoma, the association between smoking and adenocarcinoma is much weaker [8]. Thus, identifying other risk factors than tobacco smoking is necessary for further prevention of lung cancer overall and certain histologic types.

Two meta-analyses of observational cohort studies suggested an inverse association between serum vitamin D and risk of lung cancer overall [9, 10]. However, conventional observational epidemiological studies have limited capability to identify causal associations due to potential bias from confounding and reverse causation [11]. Although well-designed prospective cohort studies can reduce the possibility of reverse causation, residual and unmeasured confounding is inevitable in observational studies. Mendelian randomization (MR) studies, however, have been suggested to be able to overcome these limitations and to help make causal inferences of modifiable risk factors on health-related outcomes, provided that the crucial assumptions are satisfied [11, 12].

Therefore, we performed a MR study using three single-nucleotide polymorphisms (SNPs) as instrumental variables for serum 25-hydroxyvitamin D [25(OH)D], the primary circulating form of vitamin D, to explore potential causal associations of serum 25(OH)D levels with
incidence of lung cancer overall and different histologic types in a population-based prospective cohort.
Material and methods

Study population and data linkage

The Nord-Trøndelag Health Study (HUNT) is a large population-based health study in Norway consisting of three separate surveys: HUNT1 (1984–1986), HUNT2 (1995–1997) and HUNT3 (2006–2008). The current study was based on data from HUNT2, in which 65227 subjects aged ≥20 years living in the county of Nord-Trøndelag participated (response rate 70%). All participants completed a general questionnaire including questions on health, lifestyle and socio-economic status. Blood samples were drawn and body weight and height were measured at a clinical examination. The HUNT Research Center received updated information about deaths of all causes and emigration of the HUNT participants from the Norwegian National Registry in which the dates of such events were recorded for all people living in Norway.

Using the unique 11-digit personal identification number of all residents in Norway, the data on HUNT2 participants were linked with data from the Cancer Registry of Norway [13]. The ICD-10 (Tenth Revision of the International Statistical Classification of Diseases and Related Health Problems) topography codes C33-C34 were used to identify incident lung cancer cases among the HUNT2 study participants. Histologic types of lung cancer were classified according to the International Classification of Diseases of Oncology (ICD-O) [14]. The participants in HUNT2 were followed from the date of participation to the date of lung cancer diagnosis, death, emigration, or end of follow-up (December 31, 2014), whichever occurred first.

We excluded subjects who reported ever cancer (n=2400) in the HUNT2 questionnaire at baseline, lung cancer cases diagnosed before the participation date (n=13) in the HUNT2 study and subjects who did not have information on genotype (n=8234), leaving 54580 subjects in the analysis cohort. Moreover, a 10% random sample (n=6613) of the HUNT2
participants was selected as a subcohort for serum 25(OH)D measurement. After further excluding individuals without serum and genotype information, 5546 individuals remained in the analysis subcohort.

**Measurement and standardization of serum 25(OH)D levels**

Serum 25(OH)D level is widely recognized as the best available proxy measure for body vitamin D status [15, 16]. Serum 25(OH)D levels were measured at the HUNT Biobank using LIAISON 25-OH Vitamin D TOTAL (DiaSorin, Saluggia, Italy), a fully automated, antibody-based, chemiluminescence assay. The detection range of the assay is 10–375 nmol/L. Because seasonal fluctuations in 25(OH)D levels were expected due to the high-latitude geographical position of Norway, a cosinor model based on month of blood draw was used to calculate season-standardized 25(OH)D level (nmol/L) that represents the annual average value of 25(OH)D for each subject [17]. The season-standardized 25(OH)D was calculated using the package cosinor (version 1.1) in R (version 3.4.2).

**Genotyping and imputation of SNPs and allele score as instrumental variables**

DNA was isolated from blood samples collected in HUNT2 and stored at the HUNT biobank. Genotyping was performed using Illumina HumanCoreExome arrays as described elsewhere [18]. Imputation was performed on samples of recent European ancestry using Minimac3 (v2.0.1, http://genome.sph.umich.edu/wiki/Minimac3) [19] from a merged reference panel constructed from the Haplotype Reference Consortium panel (release version 1.1) [20] and a local reference panel based on 2201 whole-genome sequenced HUNT participants [21]. In total 3 SNPs located in or near genes for vitamin D synthesis and metabolism were selected as instrumental variables for serum 25(OH)D based on two widely-cited GWAS studies [22, 23]: rs2282679 (GC), rs12785878 (NADSYN1/DHCR7), rs10741657 (CYP2RI). Information on rs6013897 that was included in Wang et al. [23] and its proxy SNPs was not available in the HUNT study, this SNP, however, showed the weakest effect on 25(OH)D level [23, 24].
The effect allele [25(OH)D increasing allele] was coded as 1 and the other allele was coded as 0 (rs2282679: T=1; rs12785878: T=1; and rs10741657: A=1). A 25(OH)D allele score, that was a sum of the number of effect alleles of rs2282679, rs12785878 and rs10741657, was generated to increase the statistical power of the analyses [25]. The \( R^2 \) values for linkage disequilibrium between these 3 SNPs were calculated [26].

**Statistical analyses**

Linear regression was applied to calculate the F-statistic and \( R^2 \) value between SNPs or the allele score and season-standardized 25(OH)D levels. Values of F-statistic greater than 10 suggest that the SNPs or allele score are valid instrumental variables [11]. Linear regression was used to estimate the associations between the allele score and continuous covariates in order to test the assumption that the instrumental variables were not associated with potential confounders for the association between serum 25(OH)D and lung cancer; logistic regression was used in corresponding analyses of binary covariates. To test if there was a causal association between serum 25(OH)D and risk of lung cancer, we used Cox proportional hazards regression to calculate hazard ratios (HRs) with 95% confidence intervals (CIs) for the incidence of lung cancer overall or histologic types in relation to the allele score. Age was used as the time scale in the models. The proportional hazards assumption was satisfied for all SNPs and the allele score. In analyses estimating the risk of a specific histologic type, all other subtypes were censored at the date of diagnosis.

To calculate Mendelian randomization estimates of serum 25(OH)D on lung cancer risk, we generated summarized data of coefficients and standard errors from linear regression of individual SNPs on season-standardized 25(OH)D levels in the subcohort (\( n=5546 \)), as well as coefficients [\( \ln(\text{hazard ratio}) \)] and standard errors from Cox regression of individual SNPs on risk of lung cancer overall or a histologic type in the cohort (\( n=54580 \)). Inverse-variance weighted (IVW) and median-based methods were used for the summarized data to calculate
MR estimates of serum 25(OH)D for lung cancer overall and histologic types [27]. An IVW estimate of the causal effect combines the ratio estimates using each genetic variant in a fixed-effect meta-analysis model [28]. To test for pleiotropy we used MR-Egger to calculate the intercept and 95% CIs [29]. Additionally, we tested for heterogeneity between SNPs using $I^2$ and Cochran’s Q statistic. To test the robustness of our findings, we performed a two-sample MR as sensitivity analysis using summarized data of SNPs–25(OH)D association derived from a previous consortium study (n≈35000) [24].

Analyses with summarized data of individual SNPs were carried out using the package MendelianRandomization (version 0.2.2) in R (version 3.4.2). All other statistical analyses were performed with Stata/SE 14.2 (College Station, TX, USA).

**Ethics**

The study was approved by the Norwegian Regional Committees for Medical and Health Research Ethics. All participants gave their informed consent on participation in HUNT, linkage to previous HUNT surveys and specific registries.
Results

During a median follow up of 18 years, a total of 676 incident lung cancer cases were diagnosed among the 54580 cohort participants. Table 1 shows the distribution of baseline characteristics in the cohort (n=54580) and subcohort (n=5546) of the HUNT2 study. In general, the distribution of baseline characteristics was similar between the cohort and subcohort. Supplementary table 1 presents the characteristics of SNPs included in the 25(OH)D allele score in the HUNT2 study. There was no evidence of departure from the Hardy-Weinberg equilibrium for the 3 SNPs. The allele frequency was in line with that of the 1000 Genomes Phase 3 data. The $R^2$ values for linkage disequilibrium between the 3 SNPs were <0.1.

F-statistics and $R^2$ values between SNPs/the allele score and season-standardized 25(OH)D levels are presented in Table 2. The SNP rs2282679 had the highest F-statistic and $R^2$ value among the 3 SNPs, showing 4.0 nmol/L increase in 25(OH)D per effect allele. The 25(OH)D allele score had a F-statistic of 197 and accounted for 3.4% of the variation of serum 25(OH)D levels. The associations between the allele score and the potential confounders are presented in Supplementary table 2. In general, taking account of multiple testing, there were no clear associations observed.

Table 3 shows that the 25(OH)D allele score was not associated with the incidence of lung cancer overall, with HR being 0.99 (95% CI 0.93 to 1.06) per allele score. There was no clear association between the allele score and risk of any histologic type of lung cancer. Based on MR estimates using either IVW method or weighted median method, there was little evidence that genetically determined season-standardized 25(OH)D was associated with risk of lung cancer overall or any histologic type (Table 4 and Figure 1). Using the IVW method, the MR estimate HR for lung cancer overall was 0.96 (95% CI 0.54 to 1.69) per 25 nmol/L increase in the genetically determined 25(OH)D level.
As shown in Table 4, $I^2$ and Cochran’s Q statistic showed no evidence for heterogeneity between the SNPs ($I^2 0.00$, 95% CI 0.00 to 0.24, $P$ value was 0.87 for lung cancer overall). The $P$ value of the intercept by MR-Egger method was 0.79 (intercept -0.03, 95% CIs -0.25 to 0.19) for lung cancer overall, suggesting no substantial pleiotropic effect of these SNPs (Table 5).

MR estimates of a 10% increase in genetically determined 25(OH)D level with risks of lung cancer and histologic types in a two-sample MR are presented in Supplementary tables 3-4 and Supplementary figure 1 as sensitivity analyses. All estimates in the two-sample MR were similar to those derived from the primary analyses.
Discussion

Main findings
In this MR analysis of a population-based prospective cohort study including 54580 subjects, we found no substantial evidence of a causal association of serum 25(OH)D level with the incidence of lung cancer overall, SCLC, adenocarcinoma or squamous cell carcinoma.

Comparison with other studies
The finding of the current MR study is inconsistent with the conclusion from two meta-analyses of observational studies [9, 10]. The results of the meta-analyses may be largely driven by the inclusion of a large cohort study showing an inverse association between 25(OH)D levels and incidence of lung cancer [30], whereas others showed no association [9, 10]. The current study is also inconsistent with results from our own observational study from the same cohort showing that lower 25(OH)D levels were associated with a lower risk of adenocarcinoma, particularly in obese individuals [31]. The present MR analysis conformed to our speculation that residual confounding by adiposity or adiposity related factors could have biased the observational results [31].

Few MR studies have explored the potential causal association between circulating vitamin D levels and lung cancer risk. Our findings are consistent with that of the study by Dimitrakopoulou et al. who used summarized data of a consortium (TRICL-ILCCO) including a large number of lung cancer cases (n=12537) [32]. Having assumed sufficient statistical power to detect moderate effects, they found no causal association between circulating vitamin D concentration and risk of lung cancer and certain histologic types (adenocarcinoma and squamous cell carcinoma) [32]. Although the conclusions are the same, our study differs from the study by Dimitrakopoulou et al. in study design. Our study investigated the association in a homogeneous population-based prospective cohort study with a long follow-up duration, while TRICL-ILCCO mainly consisted of case-control
studies from different geographical areas and ethnicities [33]. Selection bias is more likely in case-control studies than in prospective cohort studies, and survivor bias in MR studies has recently been discussed as a methodological issue [34]. Nevertheless, we need to note that the causal effect estimates in MR studies generally reflect a life-time risk, regardless of the follow-up time.

**Strengths and limitations**

The current study is one of the first MR analyses using a long-term prospective population-based study to investigate the associations of serum 25(OH)D levels with the risk of lung cancer and histologic types. Information about diagnosis of lung cancer at the Cancer Registry of Norway is nearly complete and reasonably accurate [35]. However, there were many cases with unknown subtypes, resulting in limited statistical power in the analyses of histologic types.

Compared with observational studies, MR studies are not vulnerable to reverse causation and unmeasured confounding when the assumptions of MR studies are satisfied. We performed both one-sample MR in which we have measured serum 25(OH)D levels in a reasonably large subcohort and two-sample MR as a sensitivity analysis. Even though two-sample MR is getting common with the access to MR Base [36], one-sample MR still has its advantages, such as testing the important assumptions of MR directly. F-statistics and $R^2$ values from the regression of SNPs/allele score on 25(OH)D levels indicated sufficient strength of the instrumental variables of the exposure in the current study. The variation of 25(OH)D explained by the 3 SNPs used in the present study was larger than that explained by 4 SNPs in the study by Vimalesswaran *et al* (3.4% vs. 1.9%) [24]. We could also investigate the associations between SNPs or the allele score with a broad range of measured and reported characteristics at baseline. Even though the instruments may still be associated with unmeasured confounding, they were not associated with important confounders such as
smoking and socio-economic status in HUNT2. The last important assumption in MR is that the instrument (SNPs or allele score) should be associated with the outcome of interest (lung cancer) only via the exposure (circulating vitamin D levels). We found no violation of this essential assumption according to MR-Egger tests, but fewer genetic instruments may have a relatively low power to detect horizontal pleiotropy [29].

This study had several potential limitations. Non-participation in HUNT2 was about 30% of the population. As participants in the HUNT studies were shown to be healthier than non-participants, our findings might differ to some degree from the true situation in the general Norwegian population [37]. The applied MR analysis was based on the assumption that the exposure-outcome relationship is linear and in dose-response [11]. We were not able to investigate the non-linear association in MR due to the lack of methods for binary outcomes [38], whereas many of the reported associations between circulating 25(OH)D levels and health outcomes were non-linear [31, 39]. The sample size of this study was likely insufficient to reveal a weak to moderate effect of vitamin D on lung cancer risk based on the wide confidence intervals of the MR estimates, but the sample size of our cohort seemed adequate to detect risk factors with large effects on lung cancer, such as smoking (Supplementary table 5). Besides, our results were consistent with the null findings of the aforementioned MR study that reported a sufficient study power of a case-control design [32]. Nevertheless, consortia consisting of data from European population-based prospective studies with long follow-up duration are warranted to further investigate the causality of vitamin D on lung cancer in MR analysis. MR studies are also called for in Asia and the Middle East where populations are reported to have lower vitamin D levels than populations from Europe [40]. In addition, results from ongoing large clinical trials [41, 42] are expected to clarify the causal association of vitamin D with cancer and other adverse outcomes, especially in individuals with low vitamin D status prior to intervention.
Conclusions

In summary, Mendelian randomization analysis indicated that serum 25(OH)D levels were not causally associated with the risk of lung cancer overall or histologic types in a population-based prospective cohort study.
Funding

YQS and this work were supported by The Norwegian Cancer Society (project ID 5769155-2015) and The Research Council of Norway “Gaveforsterkning”. BMB was supported by a Research grant (#46055500-10) from The Liaison Committee for education, research and innovation in Central Norway. The K.G. Jebsen Center for Genetic Epidemiology is financed by Stiftelsen Kristian Gerhard Jebsen, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU) and Central Norway Regional Health Authority.

Disclaimer

The study has used data from the Cancer Registry of Norway. The interpretation and reporting of these data are the sole responsibility of the authors, and no endorsement by the Cancer Registry of Norway is intended nor should be inferred.

Acknowledgments

The Nord-Trøndelag Health Study (HUNT) is a collaboration between the HUNT Research Centre (Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology), the Nord-Trøndelag County Council, and the Norwegian Institute of Public Health. The authors especially thank the HUNT Research Centre laboratory personnel for the measurement of serum 25(OH)D levels.

Contributors

YQS, YC and XMM contributed to the study design. XMM contributed to data collection.

YQS, BMB, CB, SJL, SB and XMM contributed to statistical analyses of MR. YQS conducted statistical analyses, interpreted results and wrote the initial draft of the manuscript.

BMB, CB, SJL, SB, FS, YC, TILN, PRR and XMM participated in the data interpretation and helped to write the final draft of the manuscript.

Competing interests
No
Reference


Table 1. Distribution of baseline characteristics in the cohort and subcohort of the HUNT2 study, 1995–1997

<table>
<thead>
<tr>
<th></th>
<th>Cohort</th>
<th>Subcohort*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of subjects</td>
<td>54580</td>
<td>5546</td>
</tr>
<tr>
<td>Age (years)</td>
<td>49.2±16.6</td>
<td>49.1±16.6</td>
</tr>
<tr>
<td>Season-standardized 25(OH)D level (nmol/L)</td>
<td>-</td>
<td>48.3±17.1</td>
</tr>
<tr>
<td>Lung cancer cases</td>
<td>676</td>
<td>77</td>
</tr>
<tr>
<td>Sex, % (women/men)</td>
<td>52.7/47.3</td>
<td>52.8/47.2</td>
</tr>
<tr>
<td>Smoking, % (never/ever/unknown)</td>
<td>42.3/55.8/1.9</td>
<td>42.5/55.4/2.1</td>
</tr>
<tr>
<td>Family history of cancer, % (no/yes)</td>
<td>74.5/25.5</td>
<td>74.1/25.9</td>
</tr>
<tr>
<td>Education (years), % (&lt;10/≥10/unknown)</td>
<td>33.6/61.8/4.6</td>
<td>32.9/62.5/4.6</td>
</tr>
<tr>
<td>Economic difficulties, % (no/yes/unknown)</td>
<td>49.8/21.8/28.3</td>
<td>51.0/21.1/27.9</td>
</tr>
<tr>
<td>Body mass index (BMI, kg/m²), % (&lt;25/≥25/unknown)</td>
<td>39.8/59.4/0.7</td>
<td>40.6/58.8/0.6</td>
</tr>
<tr>
<td>Physical activity, % (inactive/active/unknown)</td>
<td>21.4/48.8/29.9</td>
<td>21.1/49.0/29.8</td>
</tr>
<tr>
<td>Alcohol consumption (times/month), % (never/≥1/unknown)</td>
<td>33.6/58.0/8.4</td>
<td>32.9/59.0/8.1</td>
</tr>
<tr>
<td>Chronic bronchitis, % (no/yes/unknown)</td>
<td>94.6/3.4/2.0</td>
<td>95.0/3.4/1.7</td>
</tr>
</tbody>
</table>

25(OH)D: 25-hydroxyvitamin D; HUNT2: The Nord-Trøndelag Health Study Survey 2

Data are given as percentage of subjects or mean ± standard deviation

*Those with measured serum 25(OH)D levels and genotype information
Table 2. Coefficient, F-statistic and $R^2$ value of linear regression between SNP/allele score and season-standardized 25(OH)D level (nmol/L) in subcohort of the HUNT2 study (n=5546)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Coefficient*</th>
<th>95% CI</th>
<th>$P$ value</th>
<th>F-statistic</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2282679</td>
<td>4.00</td>
<td>(3.30 to 4.71)</td>
<td>$1.63 \times 10^{-28}$</td>
<td>124</td>
<td>0.022</td>
</tr>
<tr>
<td>rs12785878</td>
<td>1.92</td>
<td>(1.26 to 2.59)</td>
<td>$1.36 \times 10^{-08}$</td>
<td>32</td>
<td>0.006</td>
</tr>
<tr>
<td>rs10741657</td>
<td>2.48</td>
<td>(1.84 to 3.13)</td>
<td>$6.98 \times 10^{-14}$</td>
<td>56</td>
<td>0.010</td>
</tr>
<tr>
<td>Allele score†</td>
<td>2.74</td>
<td>(2.36 to 3.12)</td>
<td>$5.15 \times 10^{-44}$</td>
<td>197</td>
<td>0.034</td>
</tr>
</tbody>
</table>

25(OH)D: 25-hydroxyvitamin D; CI: confidence interval; HUNT2: The Nord-Trøndelag Health Study Survey 2; SNP: Single-Nucleotide Polymorphism

*Coefficient: change of season-standardized 25(OH)D (nmol/L) per effect allele or per allele score

†The 25(OH)D allele score is sum of the number of effect alleles of rs2282679, rs12785878 and rs10741657
Table 3. The associations between the 25(OH)D allele score and risk of lung cancer overall and histologic types in the HUNT2 study (n=54580)

<table>
<thead>
<tr>
<th>Number of cases</th>
<th>Rate (100000 PY*)</th>
<th>HR†</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung cancer overall</td>
<td>676</td>
<td>73.8</td>
<td>0.99 (0.93 to 1.06)</td>
<td>0.83</td>
</tr>
<tr>
<td>SCLC</td>
<td>90</td>
<td>9.8</td>
<td>0.92 (0.77 to 1.10)</td>
<td>0.35</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>189</td>
<td>20.6</td>
<td>0.92 (0.82 to 1.04)</td>
<td>0.20</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>141</td>
<td>15.4</td>
<td>0.98 (0.85 to 1.14)</td>
<td>0.82</td>
</tr>
<tr>
<td>Other/unknown subtypes</td>
<td>256#</td>
<td>27.9</td>
<td>1.08 (0.97 to 1.21)</td>
<td>0.14</td>
</tr>
</tbody>
</table>

25(OH)D: 25-hydroxyvitamin D; CI: confidence interval; HR: hazard ratio; HUNT2: The Nord-Trøndelag Health Study Survey 2; PY: person-years; SCLC: small cell lung cancer

*TTime at risk for outcomes: 916480 person-years
†Per vitamin D-increasing allele score
#Among which 87 cases were other subtypes
<table>
<thead>
<tr>
<th>Outcome</th>
<th>MR estimate HR (95% CI)</th>
<th>IVW method</th>
<th>Weighted median method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung cancer overall</td>
<td>0.96 (0.54 to 1.69)</td>
<td>0.88</td>
<td>0.99</td>
</tr>
<tr>
<td>SCLC</td>
<td>0.58 (0.12 to 2.69)</td>
<td>0.48</td>
<td>0.82</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>0.54 (0.18 to 1.57)</td>
<td>0.26</td>
<td>0.41</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>0.64 (0.19 to 2.17)</td>
<td>0.47</td>
<td>0.52</td>
</tr>
<tr>
<td>Other/unknown subtypes</td>
<td>2.22 (0.86 to 5.75)</td>
<td>0.10</td>
<td>0.07</td>
</tr>
</tbody>
</table>

25(OH)D: 25-hydroxyvitamin D; CI: confidence interval; HR: hazard ratio; IVW: inverse-variance weighted; MR: Mendelian randomization; SCLC: small cell lung cancer
Table 5. MR-Egger pleiotropy test of associations between a 25 nmol/L increase in genetically determined season-standardized 25(OH)D and risk of lung cancer overall and histologic types in the HUNT2 study (n=54580)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>MR-Egger method</th>
<th>Intercept (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung cancer overall</td>
<td>-0.03</td>
<td>(-0.25 to 0.19)</td>
<td>0.79</td>
</tr>
<tr>
<td>SCLC</td>
<td>-0.27</td>
<td>(-1.27 to 0.72)</td>
<td>0.59</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>-0.14</td>
<td>(-0.55 to 0.27)</td>
<td>0.50</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>0.36</td>
<td>(-0.12 to 0.84)</td>
<td>0.14</td>
</tr>
<tr>
<td>Other/unknown subtypes</td>
<td>-0.09</td>
<td>(-0.45 to 0.28)</td>
<td>0.64</td>
</tr>
</tbody>
</table>

25(OH)D: 25-hydroxyvitamin D; CI: confidence interval; MR: Mendelian randomization; SCLC: small cell lung cancer