**This is Author’s Original Version (“preprint” version)**

This article has been accepted for publication in European Journal of Epidemiology published by Oxford University Press.

doi: 10.1007/s10654-017-0324-1

[**https://link.springer.com/article/10.1007%2Fs10654-017-0324-1**](https://link.springer.com/article/10.1007/s10654-017-0324-1)

**Associations of serum 25-hydroxyvitamin D level with incidence of lung cancer and histologic types in Norwegian adults – a case-cohort analysis of The HUNT Study**

Yi-Qian Sun1, \*, Arnulf Langhammer2, Chunsen Wu3, 4, Frank Skorpen1, Yue Chen5, Tom IL Nilsen2, Pål Richard Romundstad2 and Xiao-Mei Mai2

1Department of Laboratory Medicine, Children’s and Women’s Health (LBK), 2Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway. 3Institute of Clinical Research, University of Southern Denmark, Odense, Denmark. 4Odense University hospital, Odense, Denmark. 5School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Canada.

# \*Corresponding author

Yi-Qian Sun, PhD

Department of Laboratory Medicine, Children’s and Women’s Health (LBK), NTNU, Norwegian University of Science and Technology, Trondheim, Norway

Tel: +47 72576121

E-mail: yi-qian.sun@ntnu.no

# Word count

Abstract: 243

Text: 3003

# Keywords

Case-cohort study, histologic types, lung cancer, pulmonary adenocarcinoma, serum 25-hydroxyvitamin D [25(OH)D], vitamin D

# Abstract

**Background:** Previous prospective studies showed inconsistent association of serum 25-hydroxyvitamin D [25(OH)D] level with lung cancer incidence. The aim of the study was to explore the associations of serum 25(OH)D levels with incidence of lung cancer overall and different histologic types.

**Methods:** We performed a population-based prospective case-cohort study including 696 incident lung cancer cases and 5804 individuals in a subcohort who participated in the second survey of the Nord-Trøndelag Health Study in Norway. Cox proportional hazards regression models counting for the case-cohort design were used to estimate hazard ratios (HRs) with 95% confidence intervals (CIs) for lung cancer overall or histologic types in relation to serum 25(OH)D levels.

**Results:** Compared with the 4th season-specific quartile of 25(OH)D (median 68 nmol/L), lower 25(OH)D levels were associated with a lower risk for lung cancer overall and adenocarcinoma in particular. The HRs for adenocarcinoma were 0.62 (95% CI 0.41 to 0.95), 0.57 (0.37 to 0.86) and 0.55 (0.36 to 0.84) for the 1st to 3rd quartiles, respectively. Associations with adenocarcinoma were somewhat stronger in the overweight/obese subjects [HRs for 1st to 3rd quartiles: 0.55 (0.30 to 1.00), 0.38 (0.21 to 0.69) and 0.50 (0.27 to 0.90)] than in the normal weight subjects [HRs for 1st to 3rd quartiles: 1.05 (0.56 to 1.95), 0.94 (0.51 to 1.73) and 0.61 (0.33 to 1.12)].

**Conclusions:** Lower serum 25(OH)D levels were associated with an approximate 40% reduced risk of pulmonary adenocarcinoma, and the association seemed stronger in overweight/obese subjects.

# Introduction

Lung cancer has been the most common cancer type for several decades worldwide. In 2012, there were an estimated 1.8 million new cases, which represented 13% of all new cases of cancer (1). Lung cancer is also the most lethal cancer, and the overall ratio of mortality to incidence is 0.87 (1). About 1.6 million people died of lung cancer worldwide in 2012 (19.4% of the total cancer mortality). The two broad histological classes of lung cancer are small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) (2). NSCLC has two further major subtypes: adenocarcinoma and squamous cell carcinoma. Smoking is strongly associated with an increased risk of SCLC and squamous cell carcinoma, whereas the association between smoking and non-small-cell adenocarcinoma is much weaker (3). Adenocarcinoma is the most common histologic type of lung cancer in most countries, and recent research shows that the occurrence of adenocarcinoma is increasing (4, 5). Therefore, exploration of other risk factors apart from tobacco smoking and a better understanding of the pathogenesis of different histologic types of lung cancer are essential for further prevention of lung cancer.

Vitamin D has been suggested to have a number of anti-carcinogenic potentials, including inducing differentiation, and inhibiting proliferation, invasiveness, angiogenesis and metastatic properties (6, 7). Vitamin D deficiency is quite common worldwide (8, 9). In our previous study of adults living in the Central Norway, the prevalence of vitamin D deficiency [defined by serum 25-hydroxyvitamin D (25(OH)D) level <50 nmol/L] was about 60% (10). Vitamin D deficiency is likely to have impact on various chronic diseases (9, 11). The majority of previous prospective studies did not observe any association between serum 25(OH)D level and lung cancer incidence (12-18), except for one study showing an inverse association (19). These studies were generally small in sample size and did not evaluate possible associations with lung cancer histologic types.

Therefore, we performed for the first time a case-cohort study aiming to explore the potential associations of serum 25(OH)D levels with incidence of lung cancer overall and histologic types. We also evaluated if sex, active smoking and body mass index (BMI) had any modifying effects on the associations of serum 25(OH)D levels with incidence of lung cancer overall and histologic types.

# Material and methods

## Study population, data linkage and study design

The Nord-Trøndelag Health Study (HUNT) is one of the largest and most comprehensive population surveys conducted in Norway (20). All inhabitants aged 20 years or older in the county of Nord-Trøndelag were invited to participate in three separate surveys: HUNT1 (1984–1986), HUNT2 (1995–1997) and HUNT3 (2006–2008). Blood and DNA samples were collected in both HUNT2 and HUNT3 and stored in a modern biobank at the HUNT Research Center. We used data and serum samples collected in HUNT2 to assure a longer follow-up duration. Briefly, in 1995–1997 approximately 93000 adults were invited to participate in HUNT2 and 65229 people participated (response rate 70%). All participants in HUNT2 were requested to complete a general questionnaire including health and lifestyle questions and social economic status. At the clinical examination body weight and height were measured by trained nurses. The HUNT Research Center also 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 for all residents in Norway, the HUNT2 population data were linked with data from the Cancer Registry of Norway (21). The ICD-10 topography codes used for registration of lung cancer were C33-C34 (22). Histologic types were classified according to International Classification of Diseases of Oncology (ICD-O) (23).

We performed a prospective case-cohort study using Prentice method and robust estimation of the variance (24, 25). Incident cases were those diagnosed with lung cancer until December 31, 2014 among the HUNT2 study population (n=841). The subcohort was a 10% random sample of the HUNT2 participants (n=6613, including 6521 non-cases and 92 cases). We then excluded subjects who reported ever cancer in the questionnaire at baseline, lung cancer cases diagnosed before the participation date in the HUNT2 study, as well as subjects without data on the vitamin D level due to lack of serum. This left 696 incident lung cancer cases and 5725 non-cases for the statistical analyses. Figure 1 depicts the inclusion and exclusion steps of the subjects in the cases and the subcohort. The study was approved by the Norwegian Regional Committees for Medical and Health Research Ethics.

## Measurement and classification of serum 25(OH)D levels

Serum 25(OH)D levels were measured at the HUNT Biobank by 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. This is a well-developed method and has been proven as a rapid, accurate and precise tool for the measurement of serum 25(OH)D levels (26). Because seasonal fluctuations in 25(OH)D levels were expected due to the high-latitude geographical position of Norway, season-specific quartiles derived from the subcohort were used to categorize 25(OH)D levels (27). The fourth quartile was used as the reference group since the median of the fourth quartile (68.0 nmol/L) corresponded well with the cut-off level of 50.0–74.9 nmol/L that was suggested as sufficient according to the National Academy of Sciences report (28).

## Information on covariates

Data of baseline variables were collected by questionnaires or clinical examination in HUNT2. These covariates were categorized as following: sex (women, men), season of blood draw (spring: March–May, summer: June–August, fall: September–November, winter: December–February), pack-years of active smoking [0 (never smokers), 0–10.0, 10.1–20.0, ≥20.1], passive smoking exposure (never, childhood only, adulthood only, both periods), family history of cancer (Is there any family member such as father, mother, siblings who reported cancer? yes/no), education (<10, 10–12, ≥13 years), economic difficulties (During the last year, has it at any time been difficult to meet the costs of food, transportation, housing and such? yes/no), BMI (<25.0, 25.0–29.9, ≥30.0 kg/m2), physical activity (inactive or very low, low, moderate, high), alcohol consumption (never, 1–4, ≥5 times/month), and chronic bronchitis (Have you had a cough with phlegm for periods of at least 3 months during each of the last two years? yes/no).

## Statistical analyses

Cox proportional hazards regression models counting for the case-cohort design (command *stcascoh* in Stata) were used to compute hazard ratios (HRs) with 95% confidence intervals (CIs) for the incidence of lung cancer overall or histologic types in relation to serum 25(OH)D levels (25). Age was used as time axis in the models. The subjects in the case-cohort were followed up from their participation date in HUNT2 study up to the date of lung cancer diagnosis, death due to other causes, emigration or end of follow-up (December 31, 2014), whichever occurred first. When one histologic type was outcome, all other types and emigration/death were regarded as censored. The proportional hazards assumption was satisfied for serum 25(OH)D levels and covariates in most circumstances. In some occasions when lung cancer overall, NSCLC, or adenocarcinoma was the outcome, sex and economic difficulties showed non-proportional hazards and we used the *tvc* option of the *stcox* command in Stata to model them.

Two multivariable models were presented. In Model 1 sex, pack-years of active smoking, passive smoking exposure, family history of cancer, education and economic difficulties were adjusted. In Model 2 we adjusted for BMI, physical activity, alcohol consumption, and chronic bronchitis in addition to the variables included in Model 1. The associations of lifestyle factors such as BMI, physical activity, and alcohol consumption with lung cancer incidence are not yet confirmed. Chronic bronchitis might be either a confounder or a mediator in the causal pathway between 25(OH)D and lung cancer, and over-adjustment was possible if the latter were the case. Therefore, results from Model 1 were presented as the main findings in this report. Furthermore, the associations of 25(OH)D with lung cancer overall and histologic types were stratified by sex, active smoking status and BMI respectively. The potential effect modifications by these variables were evaluated by Wald tests. To reduce possible reverse causality by existing but undiagnosed lung cancer on the serum 25(OH)D levels, we performed analyses excluding the first five years of follow-up.

For sensitivity analyses, we used both the ordinary quartiles derived from the subcohort and the commonly used cut-off groups (<25.0, 25.0–49.9, 50.0–74.9, ≥75.0 nmol/L) to categorize the levels of serum 25(OH)D and controlled for season of blood draw. The fourth quartile (≥58.1 nmol/L) and the level 50.0–74.9 nmol/L were used as the reference groups respectively (28). All statistical analyses were performed with Stata/SE 14.2 (College Station, TX, USA).

# Results

Table 1 presents the distributions of baseline characteristics in the incident lung cancer cases (n=696) and non-cases from the subcohort (n=5725). In general, those with lung cancer were older and had slightly higher 25(OH)D levels than the non-cases. The distributions of other baseline characteristics differed considerably between the lung cancer cases and non-cases apart from season of blood draw and BMI. The distributions of baseline characteristics by season-specific quartiles of 25(OH)D in the subcohort are presented in Supplementary table 1.

Table 2 presents the HRs and 95% CIs for lung cancer overall and histologic types in association with serum 25(OH)D levels. Compared with the 4th season-specific quartile of 25(OH)D, lower levels of 25(OH)D were associated with a lower risk for lung cancer overall, as well as for NSCLC and adenocarcinoma. Adjusted HRs for adenocarcinoma were 0.62 (95% CI 0.41 to 0.95), 0.57 (0.37 to 0.86) and 0.55 (0.36 to 0.84) for the 1st to 3rd season-specific quartiles, respectively. Further adjustment for BMI, physical activity, alcohol consumption, and chronic bronchitis slightly attenuated these associations. However, we found no associations between serum 25(OH)D levels and risk of SCLC or squamous cell carcinoma.

We also investigated possible effect modifications by sex, active smoking and BMI on the association between serum 25(OH)D and lung cancer. Lower 25(OH)D levels were associated with reduced risks of adenocarcinoma in both women and men, with slightly stronger associations in men than in women (Table 3). Nevertheless, no statistically significant effect modification by sex was observed for lung cancer overall or for different histologic types (*p* for interaction tests >0.27 for all outcomes). There was also no evidence of effect modification by active smoking (Supplementary table 2). Analyses stratified by BMI seemed to show stronger associations of serum 25(OH)D levels with adenocarcinoma risk in overweight/obese subjects than in people who were normal weight (*p* for interaction test=0.13) (Table 4). Compared with the 4th season-specific quartile, the 1st to 3rd season-specific 25(OH)D quartiles had HRs 0.55 (95% CI 0.30 to 1.00), 0.38 (0.21 to 0.69) and 0.50 (0.27 to 0.90) respectively in overweight/obesity, whereas the corresponding estimates in normal weight people were 1.05 (0.56 to 1.95), 0.94 (0.51 to 1.73) and 0.61(0.33 to 1.12).

Finally, to address potential reverse association we performed the main analyses after excluding the first five years of follow-up (Table 5). The 1st to 3rd season-specific 25(OH)D quartiles were associated with a 48% (95% CI 0.32 to 0.85), 44% (0.35 to 0.88) and 39% (0.39 to 0.95) lower risk of pulmonary adenocarcinoma compared with the 4th quartile. Moreover, sensitivity analyses based on ordinary quartiles or commonly used cut-off values of the 25(OH)D categories produced similar results as those presented above based on the season-specific quartiles (Supplementary tables 3–5).

# Discussion

## Main findings

In this case-cohort study with 696 incident lung cancer cases, we found that the 1st to 3rd season-specific quartiles of serum 25(OH)D levels were associated with an about 40% lower risk for pulmonary adenocarcinoma compared with the highest quartile. The association between lower serum 25(OH)D levels and a lower risk of adenocarcinoma appeared stronger among the overweight/obese subjects than in the normal weight subjects, although estimates from these stratified analyses had low precision due to limited number of cases.

## Comparison with other studies

Our observation that lower serum 25(OH)D levels were associated with a lower risk of lung cancer overall and in particular pulmonary adenocarcinoma is different from the null association observed in most previous observational studies (12-18). Conversely, Afzal *et al.* showed that a 50% reduction of plasma 25(OH)D level had 19% increased risk for lung cancer overall (19). Another prospective study demonstrated that women and young participants with a higher level of vitamin D had a lower risk of lung cancer (13). However, previous studies in general had much smaller number of lung cancer cases and lacked information on histologic types.

Our findings were unexpected and contrary to the predominating hypothesis generated from *in vitro* experiments and animal models in which vitamin D was beneficial in protecting against development of cancer (7). Epidemiological studies have also shown that low levels of vitamin D was associated with increased incidence of digestive system cancers (12, 29). However, the pathophysiological roles of vitamin D in the development of cancers are complicated. As evidence, several prospective studies reported a positive relationship between vitamin D level and risk of prostate cancer (18, 30). Worthy of note, our study generally suggested a reduced risk of pulmonary adenocarcinoma in relation to lower levels of serum 25(OH)D, which cannot simply be extrapolated to an elevated risk of adenocarcinoma associated with higher levels of 25(OH)D. This was because we could not adequately evaluate the relation of higher levels of 25(OH)D with risk of lung cancer due to small number of individuals with 25(OH)D levels ≥100 nmol/L (n=56).

## Possible mechanisms and implications

Without any experimental evidence, it is difficult to speculate how lower 25(OH)D status might protect the development of adenocarcinoma of the lung, especially given that exposure to high concentrations of 1,25-dihydroxyvitamin D [1,25(OH)2D] inhibits cell proliferation and inflammation, upregulates apoptotic pathways, and inhibits angiogenesis in cell lines (6, 7). The association between lower 25(OH)D levels and reduced risk of pulmonary adenocarcinoma appeared stronger in overweight/obese subjects in our study. Observational studies have shown an inverse relation between BMI and lung cancer risk, especially among former and current smokers (31, 32). Although residual confounding by smoking might explain the inverse association of obesity with lung cancer, a recent Mendelian randomization study suggested a possible causal association between genetically determined higher BMI and a reduced risk for pulmonary adenocarcinoma (33). Higher BMI was also related to lower 25(OH)D levels, and a Mendelian randomization analysis of multiple cohorts suggested that higher BMI resulted in lower serum 25(OH)D levels (34). Therefore, vitamin D might partially explain the association between increased BMI and reduced risk of adenocarcinoma of the lung. Unfortunately, we could not adequately test this hypothesis due to the concurrent measurements of serum 25(OH)D levels and BMI at baseline. It is also highly speculative that obesity and low vitamin D status might biologically interact with each other to have a protective impact on the development of adenocarcinoma. Further research is necessary to elucidate the underlying mechanisms.

## Strengths and limitations

Our study is the first prospective case-cohort study to investigate the associations of serum 25(OH)D levels with the risk of lung cancer histologic types, including a large number of cases. Information about diagnosis of lung cancer and histologic types was recorded and updated one year after diagnosis at the Cancer Registry of Norway (21). The information was almost complete and reasonably accurate, and therefore misclassification of lung cancer was less likely (35). Serum 25(OH)D level is widely recognized as the best available approach to determine vitamin D status (36). Any misclassification of vitamin D due to measurement error would be non-differential as blood samples had been collected before the events occurred. Besides, serum samples from the cases and subcohort were placed in a random manner during the measurement so that major batch effects can be excluded. In addition, we were able to adjust for a broad panel of confounding factors in the regression models, and the season-specific quartiles, ordinary quartiles and cut-off groups showed consistent and robust associations with incidence of lung cancer overall and adenocarcinoma.

This study had several potential limitations. As participants in the HUNT studies were shown to be healthier than non-participants, our findings might differ to some degree from the factual situation in the general Norwegian population (37). Liaison immunoassay method tends to underestimate the true 25(OH)D levels (38, 39). Caution is thus warranted when our results are compared with those using other assay methods (38). In addition, our findings were derived from a homogeneous Norwegian population, which may limit the generalizability of our results. It would be important to investigate this association in other populations, such as in Asia where there is very high incidence of lung cancer, and in particular adenocarcinoma (5). Mendelian randomization studies are also called for to test possible causal association between vitamin D levels and lung cancer incidence.

Our observed reduced risk of pulmonary adenocarcinoma in association with lower 25(OH)D levels cannot be explained by the potential residual confounding of tobacco smoking. If the residual confounding of smoking could completely be removed, we would expect to see the association estimates to be more away from the null (HR=1). This is because tobacco smoking is inversely associated with serum 25(OH)D levels (10) but positively associated with lung cancer risk. However, confounding by unknown factors cannot be ruled out. Reverse association between 25(OH)D and pulmonary adenocarcinoma in our findings was less likely since the association remained after we excluded the first five years of follow-up. Moreover, 24-hydroxylase, encoded by *CYP24A1*, was overexpressed in adenocarcinoma of lung cancer (40). The 24-hydroxylase catabolizes both 25(OH)D and 1,25(OH)2D (40), which would lead to lower 25(OH)D levels in the subjects with adenocarcinoma. This could further exclude reverse causation as a possible explanation for our findings.

## Conclusions

In a prospective case-cohort study, we found that lower serum 25(OH)D levels were associated with a reduced risk of pulmonary adenocarcinoma, and the association appeared stronger in overweight/obese subjects. The unexpected findings suggest a complex role of vitamin D in cancer development in humans. Further studies are warranted to confirm or dispute our observations.

# Funding

This work was supported by The Norwegian Cancer Society (project ID 5769155-2015) and The Research Council of Norway “Gaveforsterkning”.

# 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. We are also grateful to Professor Margaret May at University of Bristol for statistical consultation.

# Contributors

YQS, AL, YC and XMM contributed to the study design. XMM and AL contributed to data collection. CW contributed to statistical analyses. YQS conducted statistical analyses, interpreted results and wrote the initial draft of the manuscript. AL, CW, FS, YC, TILN, PRR and XMM participated in the data interpretation and helped to write the final draft of the manuscript.

# Competing interests

YQS reports grants from The Norwegian Cancer Society during the conduct of the study. All other authors declare that they have no conflict of interest.

# Reference

1. GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. International Agency for Research on Cancer. 2016. <http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx>.

2. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart.: Lyon, France: IARC Press; 2004.

3. Lee PN, Forey BA, Coombs KJ. Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. BMC Cancer. 2012;12:385. doi:10.1186/1471-2407-12-385

4. Lewis DR, Check DP, Caporaso NE, Travis WD, Devesa SS. US lung cancer trends by histologic type. Cancer. 2014;120(18):2883-92. doi:10.1002/cncr.28749

5. Cheng TY, Cramb SM, Baade PD, Youlden DR, Nwogu C, Reid ME. The International Epidemiology of Lung Cancer: Latest Trends, Disparities, and Tumor Characteristics. J. Thorac. Oncol. 2016;11(10):1653-71. doi:10.1016/j.jtho.2016.05.021

6. Deeb KK, Trump DL, Johnson CS. Vitamin D signalling pathways in cancer: potential for anticancer therapeutics. Nat. Rev. Cancer. 2007;7(9):684-700. doi:10.1038/nrc2196

7. Feldman D, Krishnan AV, Swami S, Giovannucci E, Feldman BJ. The role of vitamin D in reducing cancer risk and progression. Nat. Rev. Cancer. 2014;14(5):342-57. doi:10.1038/nrc3691

8. Arabi A, El Rassi R, El-Hajj Fuleihan G. Hypovitaminosis D in developing countries-prevalence, risk factors and outcomes. Nat. Rev. Endocrinol. 2010;6(10):550-61. doi:10.1038/nrendo.2010.146

9. Palacios C, Gonzalez L. Is vitamin D deficiency a major global public health problem? J. Steroid Biochem. Mol. Biol. 2014;144 Pt A:138-45. doi:10.1016/j.jsbmb.2013.11.003

10. Sun YQ, Langhammer A, Skorpen F, Chen Y, Mai XM. Serum 25-hydroxyvitamin D level, chronic diseases and all-cause mortality in a population-based prospective cohort: the HUNT Study, Norway. BMJ Open. 2017;In press.

11. Chowdhury R, Kunutsor S, Vitezova A, et al. Vitamin D and risk of cause specific death: systematic review and meta-analysis of observational cohort and randomised intervention studies. BMJ (Clinical research ed.). 2014;348:g1903. doi:10.1136/bmj.g1903

12. Giovannucci E, Liu Y, Rimm EB, et al. Prospective study of predictors of vitamin D status and cancer incidence and mortality in men. J. Natl. Cancer Inst. 2006;98(7):451-9. doi:10.1093/jnci/djj101

13. Kilkkinen A, Knekt P, Heliovaara M, et al. Vitamin D status and the risk of lung cancer: a cohort study in Finland. Cancer Epidemiol. Biomarkers Prev. 2008;17(11):3274-8. doi:10.1158/1055-9965.EPI-08-0199

14. Ordonez-Mena JM, Schottker B, Fedirko V, et al. Pre-diagnostic vitamin D concentrations and cancer risks in older individuals: an analysis of cohorts participating in the CHANCES consortium. Eur. J. Epidemiol. 2016;31(3):311-23. doi:10.1007/s10654-015-0040-7 [doi];10.1007/s10654-015-0040-7 [pii]

15. Ordonez-Mena JM, Schottker B, Haug U, et al. Serum 25-hydroxyvitamin d and cancer risk in older adults: results from a large German prospective cohort study. Cancer Epidemiol. Biomarkers Prev. 2013;22(5):905-16. doi:10.1158/1055-9965.EPI-12-1332

16. Skaaby T, Husemoen LL, Thuesen BH, et al. Prospective population-based study of the association between serum 25-hydroxyvitamin-D levels and the incidence of specific types of cancer. Cancer Epidemiol. Biomarkers Prev. 2014;23(7):1220-9. doi:10.1158/1055-9965.EPI-14-0007

17. Weinstein SJ, Yu K, Horst RL, Parisi D, Virtamo J, Albanes D. Serum 25-hydroxyvitamin D and risk of lung cancer in male smokers: a nested case-control study. PLoS One. 2011;6(6):e20796. doi:10.1371/journal.pone.0020796

18. Wong YY, Hyde Z, McCaul KA, et al. In older men, lower plasma 25-hydroxyvitamin D is associated with reduced incidence of prostate, but not colorectal or lung cancer. PLoS One. 2014;9(6):e99954. doi:10.1371/journal.pone.0099954

19. Afzal S, Bojesen SE, Nordestgaard BG. Low plasma 25-hydroxyvitamin D and risk of tobacco-related cancer. Clin. Chem. 2013;59(5):771-80. doi:10.1373/clinchem.2012.201939

20. Krokstad S, Langhammer A, Hveem K, et al. Cohort Profile: the HUNT Study, Norway. International journal of epidemiology. 2013;42(4):968-77. doi:10.1093/ije/dys095

21. Cancer Registry of Norway. 2017. <https://www.kreftregisteret.no/>.

22. Cancer in Norway 2014 - Cancer incidence, mortality, survival and prevalence in Norway. Oslo: Cancer Registry of Norway2015.

23. International classification of diseases for oncology (ICD-O). 3rd ed: World Health Organization; 2013.

24. Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. J. Clin. Epidemiol. 1999;52(12):1165-72.

25. Coviello M. Ordinary case–cohort design and analysis. Stata Technical Bulletin STB-59. 2001.

26. Wagner D, Hanwell HE, Vieth R. An evaluation of automated methods for measurement of serum 25-hydroxyvitamin D. Clin. Biochem. 2009;42(15):1549-56. doi:10.1016/j.clinbiochem.2009.07.013

27. Wang Y, Jacobs EJ, McCullough ML, et al. Comparing methods for accounting for seasonal variability in a biomarker when only a single sample is available: insights from simulations based on serum 25-hydroxyvitamin d. Am. J. Epidemiol. 2009;170(1):88-94. doi:10.1093/aje/kwp086

28. Ross AC, Taylor CL, Yaktine AL, Del Valle HB. Dietary Reference Intakes for Calcium and Vitamin D. Washington, DC: The National Academies Press; 2011.

29. Jacobs ET, Kohler LN, Kunihiro AG, Jurutka PW. Vitamin D and Colorectal, Breast, and Prostate Cancers: A Review of the Epidemiological Evidence. J. Cancer. 2016;7(3):232-40. doi:10.7150/jca.13403

30. Meyer HE, Robsahm TE, Bjorge T, Brustad M, Blomhoff R. Vitamin D, season, and risk of prostate cancer: a nested case-control study within Norwegian health studies. Am. J. Clin. Nutr. 2013;97(1):147-54. doi:ajcn.112.039222 [pii];10.3945/ajcn.112.039222 [doi]

31. Duan P, Hu C, Quan C, et al. Body mass index and risk of lung cancer: Systematic review and dose-response meta-analysis. Sci. Rep. 2015;5:16938. doi:10.1038/srep16938

32. Yang Y, Dong J, Sun K, et al. Obesity and incidence of lung cancer: a meta-analysis. Int. J. Cancer. 2013;132(5):1162-9. doi:10.1002/ijc.27719

33. Carreras-Torres R, Haycock PC, Relton CL, 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. doi:10.1038/srep31121

34. Vimaleswaran KS, Berry DJ, Lu C, et al. Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts. PLoS. Med. 2013;10(2):e1001383. doi:10.1371/journal.pmed.1001383 [doi];PMEDICINE-D-12-01552 [pii]

35. Larsen IK, Smastuen M, Johannesen TB, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. Eur. J. Cancer. 2009;45(7):1218-31. doi:10.1016/j.ejca.2008.10.037

36. Holick MF. Vitamin D deficiency. N. Engl. J. Med. 2007;357(3):266-81. doi:10.1056/NEJMra070553

37. Langhammer A, Krokstad S, Romundstad P, Heggland J, Holmen J. The HUNT study: participation is associated with survival and depends on socioeconomic status, diseases and symptoms. BMC Med. Res. Methodol. 2012;12:143. doi:10.1186/1471-2288-12-143

38. Gaksch M, Jorde R, Grimnes G, et al. Vitamin D and mortality: Individual participant data meta-analysis of standardized 25-hydroxyvitamin D in 26916 individuals from a European consortium. PLoS One. 2017;12(2):e0170791. doi:10.1371/journal.pone.0170791

39. Schottker B, Jansen EH, Haug U, Schomburg L, Kohrle J, Brenner H. Standardization of misleading immunoassay based 25-hydroxyvitamin D levels with liquid chromatography tandem-mass spectrometry in a large cohort study. PLoS One. 2012;7(11):e48774. doi:10.1371/journal.pone.0048774

40. Chen G, Kim SH, King AN, et al. CYP24A1 is an independent prognostic marker of survival in patients with lung adenocarcinoma. Clin. Cancer Res. 2011;17(4):817-26. doi:10.1158/1078-0432.ccr-10-1789

**Table 1.** Distributions of baseline characteristics in incident lung cancer cases and non-cases from a subcohort in a case-cohort design of the HUNT2 study, 1995–1997

|  |  |  |
| --- | --- | --- |
|  | Lung cancer cases | Non-cases from a subcohort  |
| No. of subjects (n=6421) | 696 | 5725 |
| Age (years) | 61.4±10.5 | 48.4±16.7 |
| 25(OH)D level (nmol/L) | 49.1±17.5 | 47.6±17.9 |
| Sex |  |  |
|  Women | 266 (38.2%) | 2993 (52.3%) |
|  Men | 430 (61.8%) | 2732 (47.7%) |
| Season of blood draw |  |  |
|  Spring | 176 (25.3%) | 1344 (23.5%) |
|  Summer | 86 (12.4%) | 720 (12.6%) |
|  Fall | 265 (38.1%) | 1973 (34.5%) |
|  Winter | 169 (24.3%) | 1688 (29.5%) |
| Pack-years of active smoking |  |  |
|  0 (never smokers) | 34 (4.9%) | 2494 (43.6%) |
|  0–10.0  | 53 (7.6%) | 1387 (24.2%) |
|  10.1–20.0 | 177 (25.4%) | 831 (14.5%) |
|  ≥20.1  | 332 (47.7%) | 551 (9.6%) |
|  Unknown | 100 (14.4%) | 462 (8.1%) |
| Passive smoking exposure |  |  |
|  Never | 48 (6.9%) | 1098 (19.2%) |
|  Childhood only | 71 (10.2%) | 1240 (21.7%) |
|  Adulthood only | 168 (24.1%) | 925 (16.2%) |
|  Both periods | 387 (55.6%) | 2318 (40.5%) |
|  Unknown | 22 (3.2%) | 144 (2.5%) |
| Family history of cancer |  |  |
|  No | 450 (64.7%) | 4259 (74.4%) |
|  Yes | 246 (35.3%) | 1466 (25.6%) |
| Education (years) |  |  |
|  <10 | 402 (57.8%) | 1849 (32.3%) |
|  10–12 | 183 (26.3%) | 1937 (33.8%) |
|  ≥13 | 74 (10.6%) | 1702 (29.7%) |
|  Unknown | 37 (5.3%) | 237 (4.1%) |
| Economic difficulties |  |  |
|  No | 308 (44.3%) | 2902 (50.7%) |
|  Yes | 129 (18.5%) | 1222 (21.3%) |
|  Unknown | 259 (37.2%) | 1601 (28.0%) |
| Body mass index (BMI, kg/m2) |  |  |
|  Normal/underweight (<25.0) | 315 (45.3%) | 2349 (41.0%) |
|  Overweight (25.0–29.9) | 280 (40.2%) | 2459 (43.0%) |
|  Obese (≥30.0) | 98 (14.1%) | 879 (15.4%) |
|  Unknown | 3 (0.4%) | 38 (0.7%) |
| Physical activity  |  |  |
|  Inactive or very low | 180 (25.9%) | 1255 (21.9%) |
|  Low | 121 (17.4%) | 1054 (18.4%) |
|  Moderate | 85 (12.2%) | 1308 (22.8%) |
|  High | 38 (5.5%) | 504 (8.8%) |
|  Unknown | 272 (39.1%) | 1604 (28.0%) |
| Alcohol consumption (times/month) |  |  |
|  Never | 248 (35.6%) | 1865 (32.6%) |
|  1–4  | 278 (39.9%) | 2768 (48.3%) |
|  ≥5 | 92 (13.2%) | 656 (11.5%) |
|  Unknown | 78 (11.2%) | 436 (7.6%) |
| Chronic bronchitis |  |  |
|  No | 599 (86.1%) | 5453 (95.2%) |
|  Yes | 66 (9.5%) | 175 (3.1%) |
|  unknown | 31 (4.5%) | 97 (1.7%) |

25(OH)D: 25-hydroxyvitamin D; data are given as number of subjects (percentage) or mean ± standard deviation

**Table 2.** The associations of 25(OH)D level with lung cancer overall and histologic types in different models

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 25(OH)D season-specific quartile: nmol/L (median) | No. of subjects | No. of cases | Crude model | Model 1 | Model 2 |
| HR | 95% CI | HR | 95% CI | HR | 95% CI |
| Lung cancer overall | 6421 | 696\* |  |  |  |  |  |  |
|  | 1st: 10.0–41.9 (28.2)  | 1590 | 155 | 0.90 | (0.71 to 1.13) | 0.77 | (0.59 to 1.00) | 0.86 | (0.65 to 1.13) |
| 2nd: 31.2–52.9 (39.9)  | 1600 | 164 | 0.83 | (0.66 to 1.05) | 0.72 | (0.55 to 0.93) | 0.80 | (0.61 to 1.04) |
| 3rd: 40.0–65.2 (51.5)  | 1612 | 180 | 0.89 | (0.71 to 1.11) | 0.80 | (0.62 to 1.03) | 0.82 | (0.63 to 1.06) |
| 4th: 52.7–238.0 (68.0)  | 1619 | 197 | 1.00 |  | 1.00 |  | 1.00 |  |
| SCLC |  | 5886 | 93 |  |  |  |  |  |  |
|  | 1st: 10.0–41.9 (28.2)  | 1481 | 27 | 1.47 | (0.83 to 2.61) | 1.13 | (0.61 to 2.10) | 1.20 | (0.63 to 2.28) |
| 2nd: 31.2–52.9 (39.9)  | 1469 | 21 | 0.98 | (0.54 to 1.80) | 0.81 | (0.43 to 1.54) | 0.88 | (0.45 to 1.71) |
| 3rd: 40.0–65.2 (51.5)  | 1474 | 23 | 1.04 | (0.58 to 1.89) | 0.89 | (0.48 to 1.65) | 0.88 | (0.48 to 1.64) |
| 4th: 52.7–238.0 (68.0)  | 1462 | 22 | 1.00 |   | 1.00 |  | 1.00 |  |
| NSCLC |  | 6177 | 426† |  |  |  |  |  |  |
|  | 1st: 10.0–41.9 (28.2)  | 1532 | 85 | 0.77 | (0.58 to 1.03) | 0.66 | (0.48 to 0.91) | 0.74 | (0.53 to 1.03) |
| 2nd: 31.2–52.9 (39.9)  | 1550 | 109 | 0.87 | (0.66 to 1.14) | 0.76 | (0.56 to 1.02) | 0.84 | (0.62 to 1.14) |
| 3rd: 40.0–65.2 (51.5)  | 1545 | 107 | 0.83 | (0.63 to 1.09) | 0.76 | (0.56 to 1.02) | 0.78 | (0.57 to 1.05) |
| 4th: 52.7–238.0 (68.0)  | 1550 | 125 | 1.00 |   | 1.00 |  | 1.00 |  |
| Adenocarcinoma | 5967 | 192 |  |  |  |  |  |  |
|  | 1st: 10.0–41.9 (28.2)  | 1492 | 43 | 0.73 | (0.50 to 1.09) | 0.62 | (0.41 to 0.95) | 0.75 | (0.48 to 1.16) |
| 2nd: 31.2–52.9 (39.9)  | 1490 | 43 | 0.65 | (0.44 to 0.97) | 0.57 | (0.37 to 0.86) | 0.64 | (0.42 to 0.98) |
| 3rd: 40.0–65.2 (51.5)  | 1489 | 41 | 0.61 | (0.41 to 0.91) | 0.55 | (0.36 to 0.84) | 0.58 | (0.38 to 0.88) |
| 4th: 52.7–238.0 (68.0)  | 1496 | 65 | 1.00 |  | 1.00 |  | 1.00 |  |
| Squamous cell carcinoma | 5927 | 137 |  |  |  |  |  |  |
|  | 1st: 10.0–41.9 (28.2)  | 1482 | 24 | 0.90 | (0.52 to 1.56) | 0.84 | (0.48 to 1.49) | 0.90 | (0.50 to 1.62) |
| 2nd: 31.2–52.9 (39.9)  | 1493 | 46 | 1.50 | (0.94 to 2.38) | 1.33 | (0.81 to 2.17) | 1.50 | (0.90 to 2.50) |
| 3rd: 40.0–65.2 (51.5)  | 1484 | 36 | 1.13 | (0.69 to 1.85) | 1.06 | (0.63 to 1.78) | 1.11 | (0.65 to 1.88) |
| 4th: 52.7–238.0 (68.0)  | 1468 | 31 | 1.00 |  | 1.00 |  | 1.00 |  |

25(OH)D: 25-hydroxyvitamin D; CI: confidence interval; HR: hazard ratio; NSCLC: non-small cell lung cancer; SCLC: small cell lung cancer

Model 1 (main results) adjusted for sex, pack-years of active smoking, passive smoking, family history of cancer, education and economic difficulties; Model 2 adjusted for body mass index (BMI), physical activity, alcohol consumption, and chronic bronchitis in addition to Model 1; Age was used as the time axis.

\*including 177 unknown types

†including 22 other types and 75 unknown NSCLC

**Table 3.** The associations of 25(OH)D level with lung cancer overall and main histologic types stratified by sex

|  |  |  |
| --- | --- | --- |
| 25(OH)D season-specific quartile: nmol/L (median) | Women | Men |
| No. of subjects | No. of cases | HR\* | 95% CI | No. of subjects | No. of cases | HR\* | 95% CI |
| Lung cancer overall | 3259 | 266 |  |  | 3162 | 430 |  |  |
|  1st: 10.0–41.9 (28.2)  | 867 | 73 | 0.73 | (0.48 to 1.10) | 723 | 82 | 0.81 | (0.57 to 1.15) |
|  2nd: 31.2–52.9 (39.9)  | 782 | 58 | 0.68 | (0.44 to 1.05) | 818 | 106 | 0.76 | (0.55 to 1.05) |
|  3rd: 40.0–65.2 (51.5)  | 801 | 72 | 0.93 | (0.62 to 1.41) | 811 | 108 | 0.73 | (0.52 to 1.01) |
|  4th: 52.7–238.0 (68.0)  | 809 | 63 | 1.00 |   | 810 | 134 | 1.00 |  |
| NSCLC |  | 3163 | 159 |  |  | 3014 | 267 |  |  |
|  1st: 10.0–41.9 (28.2)  | 845 | 46 | 0.82 | (0.50 to 1.34) | 687 | 39 | 0.56 | (0.36 to 0.87) |
|  2nd: 31.2–52.9 (39.9)  | 762 | 35 | 0.74 | (0.44 to 1.25) | 788 | 74 | 0.79 | (0.54 to 1.13) |
|  3rd: 40.0–65.2 (51.5)  | 772 | 41 | 0.93 | (0.56 to 1.54) | 773 | 66 | 0.67 | (0.46 to 0.98) |
|  4th: 52.7–238.0 (68.0)  | 784 | 37 | 1.00 |  | 766 | 88 | 1.00 |  |
| Adenocarcinoma | 3094 | 81 |  |  | 2873 | 111 |  |  |
|  1st: 10.0–41.9 (28.2)  | 824 | 24 | 0.72 | (0.38 to 1.33) | 668 | 19 | 0.54 | (0.29 to 0.99) |
|  2nd: 31.2–52.9 (39.9)  | 746 | 15 | 0.53 | (0.26 to 1.09) | 744 | 28 | 0.58 | (0.34 to 0.98) |
|  3rd: 40.0–65.2 (51.5)  | 754 | 20 | 0.77 | (0.40 to 1.47) | 735 | 21 | 0.42 | (0.24 to 0.73) |
|  4th: 52.7–238.0 (68.0)  | 770 | 22 | 1.00 |  | 726 | 43 | 1.00 |  |

25(OH)D: 25-hydroxyvitamin D; CI: confidence interval; HR: hazard ratio; NSCLC: non-small cell lung cancer

\*Model 1 adjusted for pack-years of active smoking, passive smoking, family history of cancer, education and economic difficulties; Age was used as the time axis.

Effect modifications by sex (*p*>0.27 for all)

**Table 4.** The associations of 25(OH)D level with lung cancer overall and main histologic types stratified by body mass index (BMI)

|  |  |  |
| --- | --- | --- |
| 25(OH)D season-specific quartile: nmol/L (median) | Normal weight | Overweight/obesity |
| No. of subjects† | No. of cases | HR\* | 95% CI | No. of subjects† | No. of cases | HR\* | 95% CI |
| Lung cancer overall | 2664 | 315 |  |  | 3716 | 378 |  |  |
|  1st: 10.0–41.9 (28.2)  | 485 | 56 | 1.15 | (0.75 to 1.77) | 1084 | 97 | 0.73 | (0.51 to 1.04) |
|  2nd: 31.2–52.9 (39.9)  | 602 | 74 | 1.03 | (0.69 to 1.55) | 991 | 89 | 0.61 | (0.43 to 0.87) |
|  3rd: 40.0–65.2 (51.5)  | 718 | 85 | 0.88 | (0.60 to 1.28) | 887 | 95 | 0.78 | (0.54 to 1.11) |
|  4th: 52.7–238.0 (68.0)  | 859 | 100 | 1.00 |  | 754 | 97 | 1.00 |  |
| NSCLC | 2554 | 195 |  |  | 3582 | 228 |  |  |
|  1st: 10.0–41.9 (28.2)  | 461 | 30 | 0.91 | (0.55 to 1.52) | 1050 | 53 | 0.62 | (0.41 to 0.95) |
|  2nd: 31.2–52.9 (39.9)  | 582 | 51 | 1.12 | (0.71 to 1.76) | 961 | 57 | 0.60 | (0.40 to 0.90) |
|  3rd: 40.0–65.2 (51.5)  | 689 | 53 | 0.92 | (0.60 to 1.42) | 849 | 54 | 0.68 | (0.44 to 1.03) |
|  4th: 52.7–238.0 (68.0)  | 822 | 61 | 1.00 |  | 722 | 64 | 1.00 |  |
| Adenocarcinoma | 2460 | 92 |  |  | 3467 | 100 |  |  |
|  1st: 10.0–41.9 (28.2)  | 450 | 19 | 1.05 | (0.56 to 1.95) | 1022 | 24 | 0.55 | (0.30 to 1.00) |
|  2nd: 31.2–52.9 (39.9)  | 555 | 23 | 0.94 | (0.51 to 1.73) | 928 | 20 | 0.38 | (0.21 to 0.69) |
|  3rd: 40.0–65.2 (51.5)  | 660 | 19 | 0.61 | (0.33 to 1.12) | 822 | 22 | 0.50 | (0.27 to 0.90) |
|  4th: 52.7–238.0 (68.0)  | 795 | 31 | 1.00 |  | 695 | 34 | 1.00 |  |

25(OH)D: 25-hydroxyvitamin D; CI: confidence interval; HR: hazard ratio; NSCLC: non-small cell lung cancer

\*Model 1 adjusted for sex, pack-years of active smoking, passive smoking, family history of cancer, education and economic difficulties; Age was used as the time axis.

Effect modifications by BMI (*p*≥0.13 for all).

†41 subjects in case-cohort design had no information on BMI

**Table 5.** The associations of 25(OH)D level with lung cancer overall and main histologic types excluding the first 5 years of follow-up

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|   |   |   |   | Crude model | Model 1 | Model 2 |
|  | 25(OH)D season-specific quartile: nmol/L (median) | No. of subjects | No. of cases | HR | 95% CI | HR | 95% CI | HR | 95% CI |
| Lung cancer overall | 6056 | 559 |  |  |  |  |  |  |
|  | 1st: 10.0–41.9 (28.2)  | 1483 | 122 | 0.89 | (0.69 to 1.15) | 0.73 | (0.55 to 0.97) | 0.81 | (0.60 to 1.10) |
| 2nd: 31.2–52.9 (39.9)  | 1503 | 127 | 0.79 | (0.61 to 1.02) | 0.67 | (0.50 to 0.89) | 0.75 | (0.56 to 1.01) |
| 3rd: 40.0–65.2 (51.5)  | 1540 | 147 | 0.88 | (0.69 to 1.12) | 0.78 | (0.60 to 1.03) | 0.80 | (0.60 to 1.06) |
| 4th: 52.7–238.0 (68.0)  | 1530 | 163 | 1.00 |  | 1.00 |  | 1.00 |  |
| NSCLC | 5862 | 347 |  |  |  |  |  |  |
|  | 1st: 10.0–41.9 (28.2)  | 1437 | 68 | 0.74 | (0.54 to 1.02) | 0.60 | (0.42 to 0.85) | 0.67 | (0.46 to 0.95) |
| 2nd: 31.2–52.9 (39.9)  | 1463 | 82 | 0.77 | (0.57 to 1.04) | 0.65 | (0.47 to 0.90) | 0.72 | (0.51 to 1.01) |
| 3rd: 40.0–65.2 (51.5)  | 1485 | 90 | 0.82 | (0.61 to 1.10) | 0.73 | (0.53 to 1.01) | 0.74 | (0.54 to 1.03) |
| 4th: 52.7–238.0 (68.0)  | 1477 | 107 | 1.00 |  | 1.00 |  | 1.00 |  |
| Adenocarcinoma | 5689 | 155 |  |  |  |  |  |  |
|  | 1st: 10.0–41.9 (28.2)  | 1401 | 30 | 0.64 | (0.41 to 1.02) | 0.52 | (0.32 to 0.85) | 0.64 | (0.39 to 1.06) |
| 2nd: 31.2–52.9 (39.9)  | 1419 | 35 | 0.65 | (0.42 to 1.01) | 0.56 | (0.35 to 0.88) | 0.64 | (0.40 to 1.03) |
| 3rd: 40.0–65.2 (51.5)  | 1440 | 37 | 0.67 | (0.44 to 1.03) | 0.61 | (0.39 to 0.95) | 0.64 | (0.41 to 1.00) |
| 4th: 52.7–238.0 (68.0)  | 1429 | 53 | 1.00 |  | 1.00 |  | 1.00 |  |

25(OH)D: 25-hydroxyvitamin D; CI: confidence interval; HR: hazard ratio; NSCLC: non-small cell lung cancer

Model 1 (main results) adjusted for sex, pack-years of active smoking, passive smoking, family history of cancer, education and economic difficulties; Model 2 adjusted for body mass index (BMI), physical activity, alcohol consumption, and chronic bronchitis in addition to Model 1; Age was used as the time axis.

# Figure legend

**Figure 1.** Inclusion of subjects to the case and subcohort groups in the HUNT2 study. The cases present in the subcohort (n=92 originally sampled; n=79 included in analyses) are duplicates of lung cancer cases in the case group.

