

Article type: Original Article

Title: No association between circulating concentrations of vitamin D and risk of lung cancer: An analysis in 20 prospective studies in the Lung Cancer Cohort Consortium (LC3)

Author List: D.C. Muller^{1,2*}, A.M. Hodge^{3,4*}, A. Fanidi^{1,5}, D. Albanes⁶, X.M. Mai⁷, X.O. Shu⁸, S.J. Weinstein⁶, T.L. Larose^{1,9}, X. Zhang¹⁰, J. Han^{11,12}, M.J. Stampfer^{10,13,14}, S.A. Smith-Warner^{13,14}, J. Ma¹⁰, J.M. Gaziano^{15,16,17}, H.D. Sesso^{13,15,16}, V.L. Stevens¹⁸, M.L. McCullough¹⁸, T.M. Layne⁶, R. Prentice¹⁹, M. Pettinger¹⁹, C.A. Thomson²⁰, W. Zheng⁸, Y.T. Gao²¹, N. Rothman⁶, Y.B. Xiang²², H. Cai²³, R. Wang²⁴, J.M. Yuan^{24,25}, W.P. Koh²⁶, L.M. Butler^{25,27}, Q. Cai²³, W.J. Blot²³, J. Wu²³, P.M. Ueland^{28,29}, Ø. Midttun³⁰, A. Langhammer³¹, K. Hveem^{9,31}, M. Johansson³², J. Hultdin³³, K. Grankvist³³, A. A. Arslan^{34,35}, L. Le Marchand³⁶, G. Severi^{3,37,38#}, M. Johansson^{1#}, P. Brennan^{1#}

* denotes co-first authors

denotes co-last authors

1. Genetic Epidemiology Group, International Agency for Research on Cancer, Lyon, France
2. Department of Epidemiology and Biostatistics, Imperial College London, London, United Kingdom
3. Cancer Epidemiology Center, Cancer Council Victoria, Melbourne, Australia
4. Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Parkville, Australia
5. MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom

6. Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, USA
7. Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology, Trondheim, Norway
8. Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, USA
9. K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health & Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway
10. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, USA
11. Department of Epidemiology, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, USA
12. Melvin and Bren Simon Cancer Center, Indiana University, Indianapolis, USA
13. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, USA
14. Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, USA
15. Division of Aging, Department of Medicine, Brigham and Women's Hospital, Boston, USA
16. Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, USA
17. Boston VA Medical Center, Boston, USA
18. Epidemiology Research Program, American Cancer Society, Atlanta, USA
19. Division of Public Health Sciences Fred Hutchinson Cancer Research Center, Seattle, USA
20. Health Promotion Sciences, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, USA
21. Department of Epidemiology, Shanghai Cancer Institute, Shanghai Jiaotong University, Shanghai, China

22. State Key Laboratory of Oncogene and Related Genes & Department of Epidemiology, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China
23. Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, USA
24. UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, USA
25. Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, USA
26. Duke-NUS Graduate Medical School Singapore, Singapore, Singapore
27. Division of Cancer Control and Population Sciences, University of Pittsburgh Cancer Institute, Pittsburgh, USA
28. Laboratory of Clinical Biochemistry, Department of Clinical Science, University of Bergen, Bergen, Norway
29. Haukeland University Hospital, Bergen, Norway
30. Bevital AS, Bergen, Norway
31. HUNT Research Centre, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology, Levanger, Norway
32. Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden
33. Department of Medical Biosciences, Clinical Chemistry, Umeå University, Umeå, Sweden
34. Departments of Obstetrics and Gynecology, Population Health and Environmental Medicine, New York University School of Medicine, New York, USA
35. Department of Population Health and Environmental Medicine, New York University School of Medicine, New York, USA
36. Epidemiology Program, Cancer Research Center of Hawaii, University of Hawaii
37. Italian Institute for Genomic Medicine (IIGM), Torino, Italy.

38. Centre de Recherche en Epidémiologie et Santé des Populations (CESP)
UMR1018 Inserm, Facultés de Médecine Université Paris-Saclay, Villejuif,
France.

Corresponding Authors:

Dr. Paul Brennan, Genetic Epidemiology Group, International Agency for
Research on Cancer (IARC-WHO), 150 cours Albert Thomas, Lyon, 69008,
France, Tel: +33 4 72 73 85 33, Email: gep@iarc.fr

AND

Dr. David C. Muller, Department of Epidemiology and Biostatistics, Imperial
College London, Norfolk Place, London W21PG, United Kingdom, Tel: +44
(0)20 7594 0856, Email: david.muller@imperial.ac.uk

Abstract

Background: There is observational evidence suggesting that high vitamin D concentrations may protect against lung cancer. To investigate this hypothesis in detail, we measured circulating vitamin D concentrations in pre-diagnostic blood from 20 cohorts participating in the Lung Cancer Cohort Consortium (LC3).

Patients and methods: The study included 5,313 lung cancer cases and 5,313 controls selected from. Blood samples for the cases were collected, on average, 5 years prior to lung cancer diagnosis. Controls were individually matched to the cases by cohort, sex, age, race/ethnicity, date of blood collection, and smoking status in 5 categories. Liquid chromatography coupled with tandem mass spectrometry was used to separately analyze 25-hydroxyvitamin D₂ (25(OH)D₂) and 25-hydroxyvitamin D₃ (25(OH)D₃) and their concentrations were combined to give an overall measure of 25(OH)D. We used conditional logistic regression to calculate odds ratios (OR) and 95% confidence intervals (CI) for 25(OH)D as both a continuous and categorical variable.

Results: Overall, no apparent association between 25(OH)D and risk of lung cancer was observed (multivariable adjusted OR for a doubling in concentration: 0.98, 95% confidence interval: 0.91, 1.06). Similarly, we found no clear evidence of interaction by cohort, sex, age, smoking status, or histology.

Conclusion: This study did not support an association between vitamin D concentrations and lung cancer risk.

Keywords: serum 25-hydroxyvitamin D, vitamin D, lung cancer, case-control, prospective, consortium

Key Message: Results from this prospective study of 20 international cohorts show no association between circulating concentrations of vitamin D and risk of lung cancer.

Introduction

Lung cancer is the most common cause of cancer death worldwide, accounting for over 20% of all cancer deaths (1). Although avoidance of tobacco consumption remains the most important strategy for lung cancer prevention, a substantial proportion of lung cancer cases occur among former smokers, especially in countries where tobacco cessation campaigns have been successful. Identifying additional strategies to reduce lung cancer risk would be of particular relevance for high risk groups such as former smokers, and could have a substantial public health impact.

There is evidence that higher circulating vitamin D concentrations may reduce the risk of developing lung cancer (2,3). Vitamin D is produced primarily in the skin after exposure to ultraviolet B (UVB) radiation from sunlight. A secondary source of vitamin D is through diet or via dietary supplements (4). Vitamin D is subsequently hydroxylated in the liver to form 25-hydroxyvitamin D (25(OH)D), the major circulating metabolite of vitamin D, which is then converted into its active form (1,25-dihydroxyvitamin D) in the kidneys and other tissues (5). The necessity of adequate concentrations of vitamin D for bone health is well known (6), and vitamin D has also been implicated in the

development of several cancers, particularly colorectal cancer (7,8). As reported in two meta-analyses, a number of prospective cohort and nested case-control studies have examined vitamin D concentrations, measured as 25-hydroxyvitamin D (25(OH)D) in blood samples, in terms of risk of lung cancer. In the first meta-analysis, pooled results from 9 prospective studies reported a 17% decrease in lung cancer risk between participants with 'high' and 'low' concentrations of 25(OH)D (2). The second meta-analysis reported similar results, although ten of the studies were included in the first meta-analysis (2). Pooled results from the second meta-analysis showed a nonlinear inverse association between 25(OH)D and lung cancer risk (3).

While these meta-analyses are informative, they are clearly constrained by the manner in which results are reported in the literature, and cannot account for variation that may occur between studies, including use of different laboratories for measurement of vitamin D. In order to provide a more definitive answer to the role of circulating vitamin D and risk of lung cancer, we conducted a pooled analysis of 25(OH)D concentrations in 5,313 case-control pairs from 20 cohorts within the Lung Cancer Cohort Consortium (LC3).

Methods

The LC3 was established in 2009 as a project within the US National Cancer Institute (NCI) Cohort Consortium as a coordinated large-scale effort to study the role of B-vitamins in lung cancer (9). Inclusion criteria for participating cohorts included the occurrence of at least 200 incident lung cancer cases

with baseline questionnaire data and either plasma or serum samples – typically cryopreserved at $<80^{\circ}\text{C}$. Twenty cohorts fulfilled those criteria and agreed to participate in the LC3, resulting in a combined cohort population of over 2,000,000 participants from Asia, Australia, Europe, and North America. All cohorts are listed in Supplementary Table 1, along with the numbers of cases and controls contributed. Brief details on design of the cohorts and their follow-up procedures are also provided in the supplementary materials.

Selection of cases and controls

Lung cancer cases were defined on the basis of the International Classification of Diseases for Oncology, Second Edition (ICD-O-2), and included all invasive cancers coded as C34.0-C34.9. Altogether, 11,399 incident lung cancer cases with pre-diagnostic blood samples were identified from the participating cohorts. We selected a total of 5,545 lung cancer cases for subsequent blood based analysis, and to optimize the statistical power in smoking stratified risk analysis, never and former smoking cases were oversampled. For each case, one control was randomly chosen from risk-sets consisting of all cohort members alive and free of cancer (except non-melanoma skin cancer) at the time of diagnosis of the index case. Matching criteria were cohort, sex, date of blood collection (± 1 month, relaxed to ± 3 months for sets without available controls), and date of birth (± 1 year, relaxed to ± 3 years for sets without available controls), as well as smoking status in 5 categories: never smokers, short and long-term quitters among former smokers (<10 years, ≥ 10 years since quitting), and light and heavy smokers among current smokers (< 15 , ≥ 15 cigarettes per day). After excluding cases

who were not correctly matched on smoking status ($n=124$ cases), who had insufficient blood samples ($n=42$), or had a revised date of diagnosis prior to blood draw ($n=13$), a total of 5,364 lung cancer case-control pairs remained eligible for analysis. Further exclusion of 51 participants whose matched pair was missing a 25(OH)D value resulted in a final set of 5,313 matched case-control pairs.

Biochemical analyses

Vitamin D analyses were conducted as part of a coordinated laboratory analysis along with a panel of 40 other biomarkers, focused primarily on B-vitamins and other biomarkers that are involved in the one carbon metabolism pathway (9). Blood samples from all cases and controls were sent on dry ice to the Bevital AS laboratory (Bergen, Norway, www.bevital.no). Liquid chromatography coupled with tandem mass spectrometry was used to separately analyze 25-hydroxyvitamin D₂ (25(OH)D₂) and 25-hydroxyvitamin D₃ (25(OH)D₃) (11). The limit of detection was 3.3 nmol/L, and within-day and between-day coefficients of variation ranged from 4.4%–8.2%. Circulating cotinine was also assessed with liquid chromatography coupled with tandem mass spectrometry. The limit of detection was 1 nmol/L, and the within-day and between-day coefficients of variation ranged from 2%–6%. The laboratory is Vitamin D External Quality Assessment Scheme–certified (DEQAS, London, United Kingdom, www.deqas.org).

Statistical Analysis

We used conditional logistic regression to calculate odds ratios (ORs) and 95% confidence intervals (CIs) for 25(OH)D as both a continuous and categorical variable. For categorical analyses, 25(OH)D was grouped according to the quartiles of its distribution among control participants, as well as an alternative grouping with extreme and moderate deficiency (<25 nmol/L) as the reference category. For continuous analyses, 25(OH)D was \log_2 transformed, so ORs correspond to the expected fold-change in odds of lung cancer for a doubling in 25(OH)D concentration. All models were adjusted for circulating cotinine in four groups, defined by quartiles of cotinine among current smoking participants. We investigated potential interactions between 25(OH)D and sex, smoking status, cohort region, age at baseline, body mass index (BMI, kg/m²), time between blood draw and lung cancer diagnosis, as well as histological subtype. These analyses were conducted using R version 3.3.2 (12).

Because 25(OH)D₃ is strongly affected by exposure to UVB radiation, we used season-adjusted concentrations. We modeled seasonal variation of log transformed 25(OH)D₃ using two pairs of sine and cosine functions of day of blood collection in a hierarchical linear regression model, allowing model parameters to vary by cohort. The use of trigonometric functions allows estimation of the periodic variation in 25(OH)D₃ and produces smooth predictions with no artificial discontinuities from season to season, or year to year. We included 2 pairs of sine and cosine functions in the models because the inclusion of additional terms did not improve model fit, nor did it substantially affect parameter estimates in the final models. For each participant, season-adjusted 25(OH)D was calculated by adding or subtracting

the expected seasonal deviation from the mean 25(OH)D₃ concentration based on the day of blood draw, and adding the 25(OH)D₂ concentration. These Bayesian hierarchical models were fitted using Stan version 2.9.1 via Rstan (13,14).

Results

Demographic and baseline characteristics of the 5,313 case-control pairs are presented in Table 1, both overall and by geographic region. Participants from the US had higher circulating concentrations of total 25(OH)D on average compared with participants from Asia, or Europe and Australia. Table 2 shows the histological subtype of lung cancer cases, as well as the distribution of time between baseline blood collection and diagnosis.

Figure 1 shows the observed 25(OH)D concentrations by day of blood draw, as well as the model-estimated mean concentrations, for each cohort separately. Figure 2 shows the season-adjusted and observed 25(OH)D concentrations. ORs for risk of lung cancer by category of 25(OH)D concentration, as well as for a doubling in concentration, are presented in Table 3. No apparent dose-response association was observed between 25(OH)D and risk of lung cancer overall (OR for a doubling in concentration: 0.98, 95% CI: 0.91, 1.06), nor for specific categories of 25(OH)D concentration. Further, we found little evidence of interaction by any demographic factor or individual participant characteristic (Figure 3). There was some evidence of an inverse association between 25(OH)D and risk of lung cancer in the European cohorts (OR for a doubling in concentration 0.75, 95% CI: 0.61, 0.92), but statistical evidence for an interaction was weak ($p =$

0.05). Cohort-specific estimates suggest that any association among European participants was driven by MDCS and NSHDC, the two Swedish cohorts), although we found no overall evidence of interaction by cohort (Figure 4). Sensitivity analyses in which we excluded circulating cotinine from the models provided similar results to the fully adjusted models (Supplementary Figure 1). Using 25(OH)D₃ only also yielded similar results (Supplementary Figure 2).

Discussion

Based on a comprehensive analysis of over 5,000 case-control pairs from 20 prospective cohort studies, we found no association between circulating vitamin D concentrations and risk of subsequent lung cancer diagnosis.

Our results contrast with the two previous meta-analyses that reported inverse associations between circulating vitamin D concentrations and lung cancer risk (2,3). As always with meta-analyses, it is possible that the collection of published studies does not represent the total sum of all conducted studies, whereby studies with null results are more likely to go unpublished. Although neither study reported evidence for publication bias, we note that the results were strongly influenced by one study from Copenhagen with approximately 10,000 participants and 507 incident lung cancer cases with 25-year follow-up (15), accounting for 65% of the association reported by Zhang et al (2). The second meta-analysis by Chen et al. was based on 10 cohorts (3), but there was considerable overlap between included studies in both the Chan and Zhang meta-analyses, and both included the aforementioned Copenhagen cohort study (15). It should therefore be no surprise that the pooled estimates

across these two meta-analyses suggest a similar inverse association with risk.

Median 25(OH)D concentrations in the Copenhagen cohort were 41 nmol/L overall, and 37 nmol/L among those who went on to develop a tobacco related cancer, similar to approximately the first fourth of 25(OH)D for the LC3 cohorts overall. It is of interest that Afzal et al. also observed significant associations with vitamin D for other tobacco related cancers including head and neck cancer, bladder and kidney cancer in the Copenhagen cohort (15). Given the limited adjustment for tobacco use in the Copenhagen cohort (i.e., pack-years only), and the strong association between vitamin D concentrations and smoking intensity, this would suggest the possibility that residual confounding by smoking may explain at least some of the observed inverse associations of vitamin D concentrations with lung cancer risk, as well as with other tobacco related cancers in that study (15). In our study, estimates were similar in models that were not adjusted for circulating cotinine, indicating that residual confounding is unlikely to fully explain these discrepant results. Further, the estimates from the Copenhagen cohort are consistent with the estimates that we present for the two cohorts from Sweden. The absence of evidence supporting an interaction by cohort, however, suggests that any observed inverse associations in our study are likely to be consistent with sampling variability.

Our results, based on over 5,000 incident lung cancer cases and 5,000 individually matched controls, clearly point to a lack of association for risk of

lung cancer across a broad range of vitamin D concentrations. Our study has several strengths including centralized biochemical analysis of pre-diagnostic vitamin D and detailed control for tobacco exposure among current smokers using cotinine concentrations, an objective measure of recent tobacco exposure. There is a lack of consensus surrounding terminology and vitamin D concentrations that are considered adequate or deficient. The 2011 US Institute of Medicine report considers individuals with concentrations below 30 nmol/L at risk of deficiency, and those below 50 nmol/L at risk of inadequacy (16). An alternative definition considers concentrations below 50 nmol/L as deficient, with further subclassifications of mild (25-50 nmol/L), moderate (12.5-25 nmol/L) and severe (<12.5 nmol/L) deficiency (17). Our reference exposure group included participants with vitamin D concentrations between 7 and 41 nmol/L, and therefore included a broad range of persons with vitamin D deficiency or inadequacy. Even when we further restricted the reference category to include only those participants with vitamin D concentrations below 25 nmol/L, we found no evidence for an association for adequate vitamin D concentrations with lower lung cancer risk.

In summary, we found no overall evidence for an association between circulating concentrations of vitamin D and risk of lung cancer in 20 prospective cohort studies from the Asia, Australia and Europe, and the US, but some suggestion of an inverse association in two Swedish cohorts. In light of this, we consider that vitamin D supplementation is unlikely to prove broadly effective for the primary prevention of lung cancer. Ongoing cancer prevention trials testing vitamin D supplements may eventually provide

additional evidence on whether or not increases in vitamin D concentrations translate to reductions in lung cancer risks or not.

Funding Acknowledgements

The Lung Cancer Cohort Consortium (LC3) was supported by Grant NIH / NCI (n° 1U01CA155340-01) and NHMRC (Grant ID: 1050198). DCM was supported by an IARC/Australia Fellowship funded by Cancer Council Australia, and a Cancer Research UK Population Research Fellowship. TLL was supported by The Research Council of Norway (grant number 267776/H10). The work of TLL presented in this paper was undertaken during a postdoctoral placement at the International Agency for Research on Cancer, within the framework of an agreement between the Research Council of Norway and the Norwegian University of Science and Technology. The funding organizations had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript.

Cohort specific acknowledgements:

The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. The authors thank the WHI investigators and staff for their dedication, and the

study participants for making the program possible. A full listing of WHI investigators can be found at:

<http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf>

SWHS was/is supported by R37 CA070867 and UM1 CA182910, SMHS, by R01 CA082729 and UM1 CA173640 from the US National Cancer Institute.

Declarations

Drs Ueland and Middtun reports that they are members of the steering board of the nonprofit Foundation to Promote Research Into Functional Vitamin B12 Deficiency. No other disclosures were reported.

We thank the study participants and LC3 consortium members for their contributions in making this study possible.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015 Mar 1;136(5):E359–86.
2. Zhang L, Wang S, Che X, Li X. Vitamin D and Lung Cancer Risk: A Comprehensive Review and Meta-Analysis. *Cell Physiol Biochem*. 2015 May 4;36(1):299–305.
3. Chen G-C, Zhang Z-L, Wan Z, Wang L, Weber P, Eggersdorfer M, et al. Circulating 25-hydroxyvitamin D and risk of lung cancer: a dose-response meta-analysis. *Cancer Causes Control CCC*. 2015 Dec;26(12):1719–28.

4. 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 May;14(5):342–57.
5. Holick MF. Resurrection of vitamin D deficiency and rickets. *J Clin Invest*. 2006 Aug;116(8):2062–72.
6. Holick MF. Vitamin D deficiency. *N Engl J Med*. 2007 Jul 19;357(3):266–81.
7. Dou R, Ng K, Giovannucci EL, Manson JE, Qian ZR, Ogino S. Vitamin D and colorectal cancer: molecular, epidemiological and clinical evidence. *Br J Nutr*. 2016 May;115(9):1643–60.
8. Ekmekcioglu C, Haluza D, Kundi M. 25-Hydroxyvitamin D Status and Risk for Colorectal Cancer and Type 2 Diabetes Mellitus: A Systematic Review and Meta-Analysis of Epidemiological Studies. *Int J Environ Res Public Health*. 2017 Jan 28;14(2):127.
9. Fanidi A, Muller DC, Yuan JM, Stevens VL, Weinstein SJ, Albanes D, Prentice R, Thomsen CA, Pettinger M et al. Circulating Folate, Vitamin B6, and Methionine in Relation to Lung Cancer Risk in the Lung Cancer Cohort Consortium (LC3). *J Natl Cancer Inst*.doi: 10.1093/jnci/djx119
10. Midttun Ø, Ueland PM. Determination of vitamins A, D and E in a small volume of human plasma by a high-throughput method based on liquid chromatography/tandem mass spectrometry. *Rapid Commun Mass Spectrom*. 2011;25(14):1942–1948.
11. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2015. Available from: <http://www.R-project.org/>
12. Stan Development Team. The Stan C++ Library, Version 2.9.0 [Internet]. 2016. Available from: <http://mc-stan.org>
13. Stan Development Team. RStan: the R interface to Stan. R package version 2.9.0 [Internet]. 2016. Available from: <http://mc-stan.org>
14. Afzal S, Bojesen SE, Nordestgaard BG. Low Plasma 25-Hydroxyvitamin D and Risk of Tobacco-Related Cancer. *Clin Chem*. 2013 May 1;59(5):771–80.
15. Institute of Medicine (US) Committee to Review Dietary Reference Intakes for Vitamin D and Calcium. Dietary Reference Intakes for Calcium and Vitamin D [Internet]. Ross AC, Taylor CL, Yaktine AL, Del Valle HB, editors. Washington (DC): National Academies Press (US); 2011 [cited 2017 Feb 28]. (The National Academies Collection: Reports funded by National Institutes of Health). Available from: <http://www.ncbi.nlm.nih.gov/books/NBK56070/>

16. Stroud ML, Stilgoe S, Stott VE, Alhabian O, Salman K. Vitamin D - a review. *Aust Fam Physician*. 2008 Dec;37(12):1002–5.

Figures captions

Figure 1: Circulating log 25(OH)D as a function of day-of-year of blood draw. Scattered points are the observed concentrations, and the lines are the predicted cohort-specific mean concentrations given the calendar day on which blood was drawn. Predictions were made by regressing log 25(OH)D concentrations on sine and cosine functions of calendar day. A hierarchical model was used to allow the model coefficients to vary between cohorts.

Figure 2: Season-adjusted log 25(OH)D versus observed concentrations. Adjusted log 25(OH)D was calculated by subtracting the seasonal component of the model predicted concentration from the observed concentration. The gradient of colour varies with the magnitude of the adjustment. This adjustment ranged from a subtraction of over 0.2 log-units (for those with blood drawn during the height of summer, magenta points) to an addition of over 0.2 log-units (for those with blood drawn in the middle of winter, dark blue points).

Figure 3: Odds ratios for a doubling in 25(OH)D concentration overall, and by demographic and individual characteristics. Estimates are from conditional logistic regression models conditioned on matched case set, and adjusted for four categories of circulating cotinine. *P*-values are from likelihood ratio tests of the interaction terms between 25(OH)D and each covariate.

Figure 4: Odds ratios for a doubling in 25(OH)D concentration separately for each cohort. Estimates are from a conditional logistic regression model conditioned on matched case set, and adjusted for four categories of

circulating cotinine. The *P*-value is from likelihood ratio test of the interaction terms between 25(OH)D cohort.

Supplementary Figure 1: Odds ratios for a doubling in 25(OH)D concentration overall, and by demographic and individual characteristics, not adjusted for circulating cotinine. Estimates are from conditional logistic regression models conditioned on matched case set. *P*-values are from likelihood ratio tests of the interaction terms between 25(OH)D and each covariate.

Supplementary Figure 2: Odds ratios for a doubling in 25(OH)D3 concentration overall, and by demographic and individual characteristics. Estimates are from conditional logistic regression models conditioned on matched case set, and adjusted for four categories of circulating cotinine. *P*-values are from likelihood ratio tests of the interaction terms between 25(OH)D3 and each covariate.

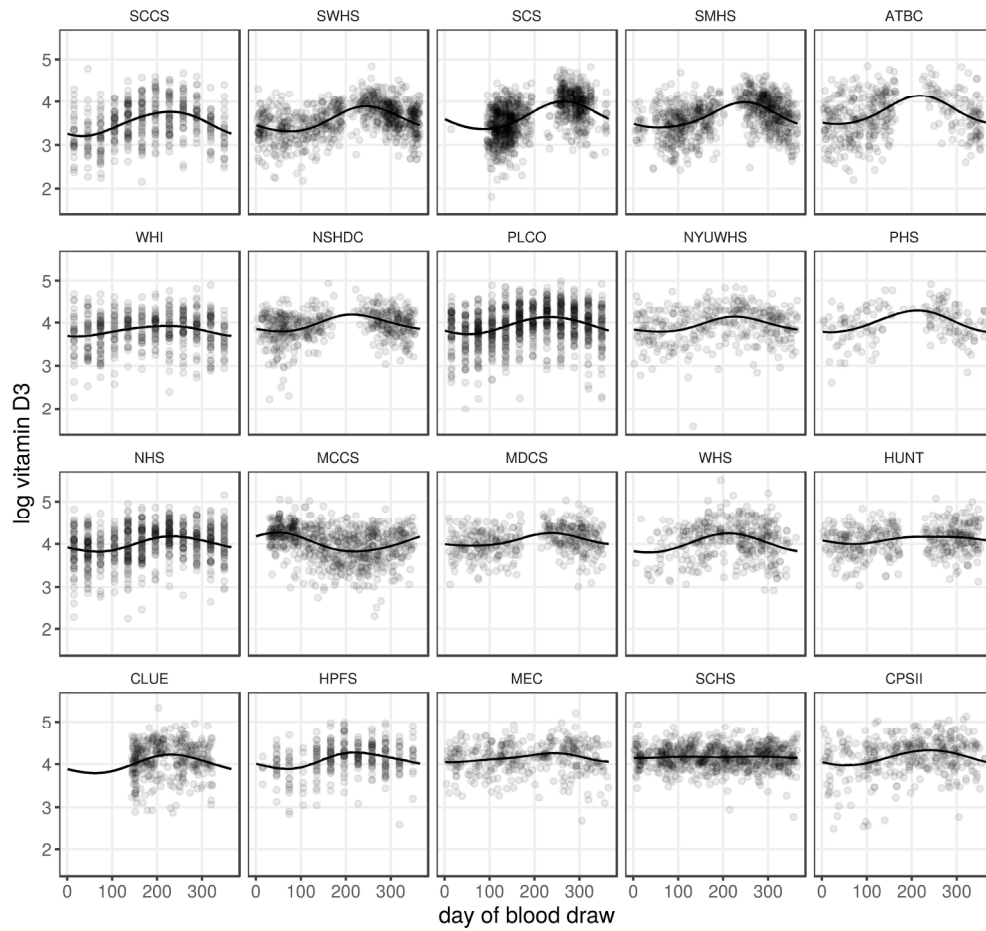


Figure 1: Circulating log 25(OH)D as a function of day-of-year of blood draw. Scattered points are the observed concentrations, and the lines are the predicted cohort-specific mean concentrations given the calendar day on which blood was drawn. Predictions were made by regressing log 25(OH)D concentrations on sine and cosine functions of calendar day. A hierarchical model was used to allow the model coefficients to vary between cohorts.

216x205mm (300 x 300 DPI)

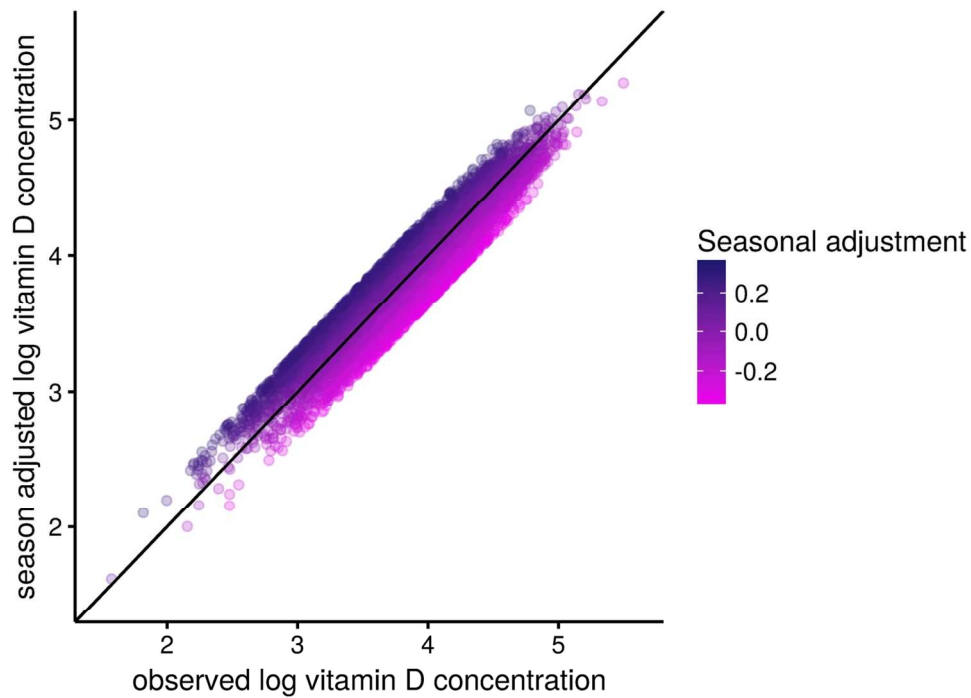


Figure 2: Season-adjusted log 25(OH)D versus observed concentrations. Adjusted log 25(OH)D was calculated by subtracting the seasonal component of the model predicted concentration from the observed concentration. The gradient of colour varies with the magnitude of the adjustment. This adjustment ranged from a subtraction of over 0.2 log-units (for those with blood drawn during the height of summer, magenta points) to an addition of over 0.2 log-units (for those with blood drawn in the middle of winter, dark blue points).

127x90mm (300 x 300 DPI)

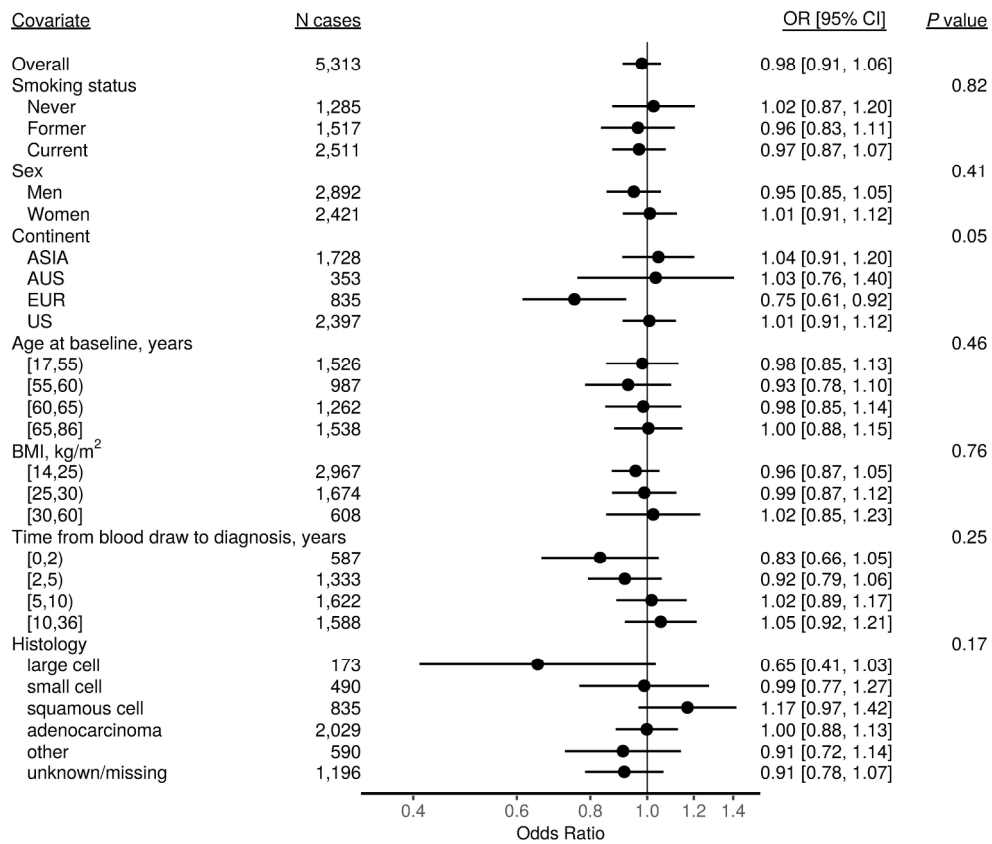


Figure 3: Odds ratios for a doubling in 25(OH)D concentration overall, and by demographic and individual characteristics. Estimates are from conditional logistic regression models conditioned on matched case set, and adjusted for four categories of circulating cotinine. *P*-values are from likelihood ratio tests of the interaction terms between 25(OH)D and each covariate.

203x180mm (300 x 300 DPI)

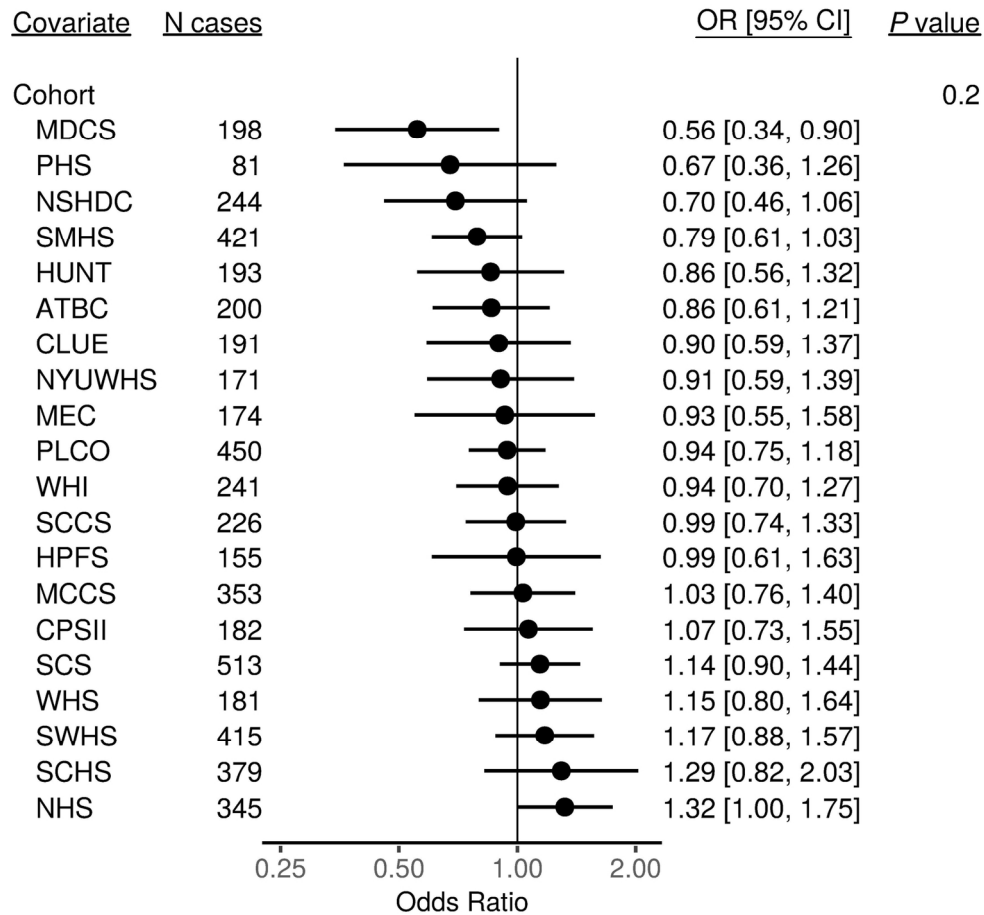
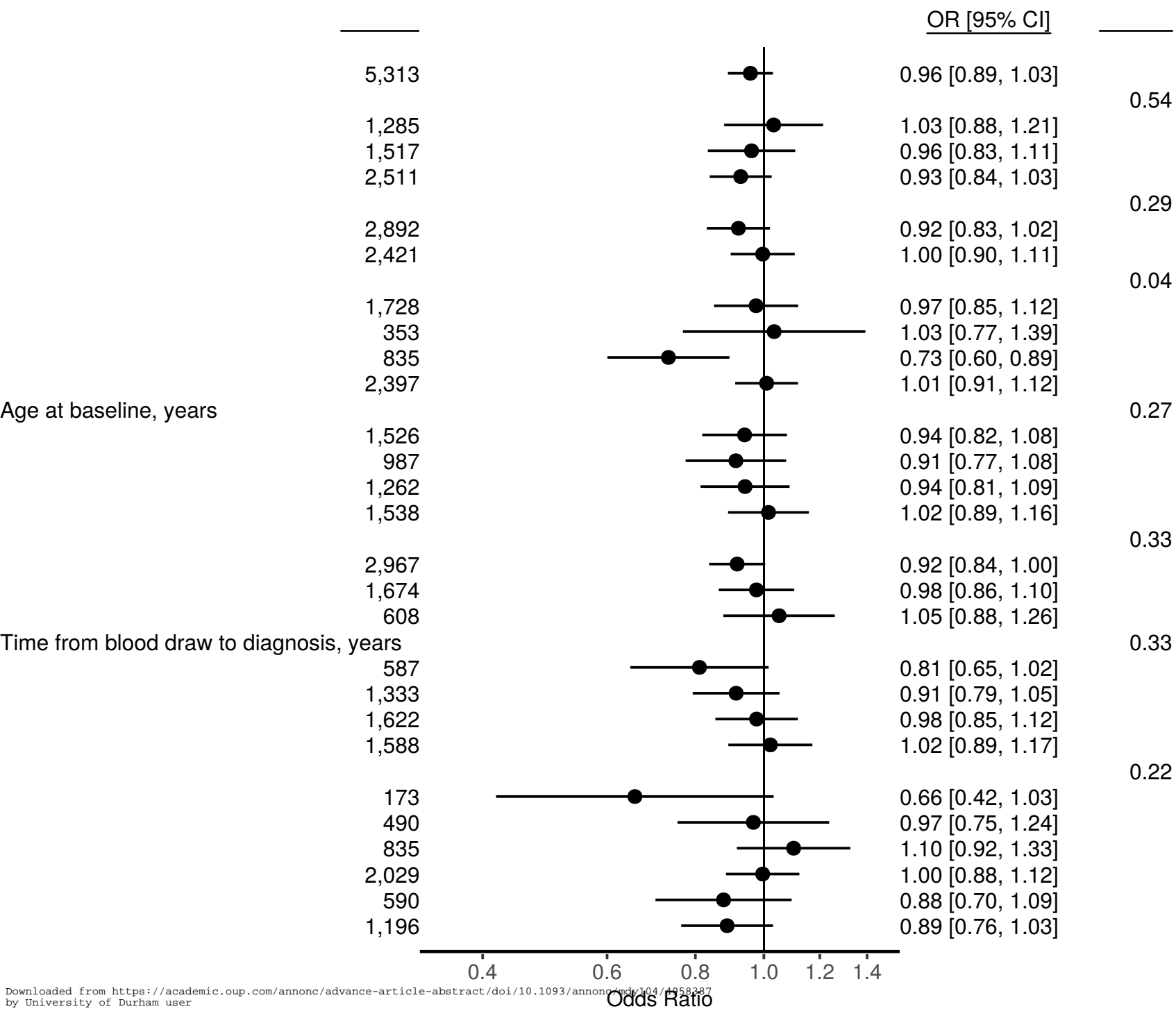


Figure 4: Odds ratios for a doubling in 25(OH)D concentration separately for each cohort. Estimates are from a conditional logistic regression model conditioned on matched case set, and adjusted for four categories of circulating cotinine. The *P*-value is from likelihood ratio test of the interaction terms between 25(OH)D cohort.

153x154mm (300 x 300 DPI)



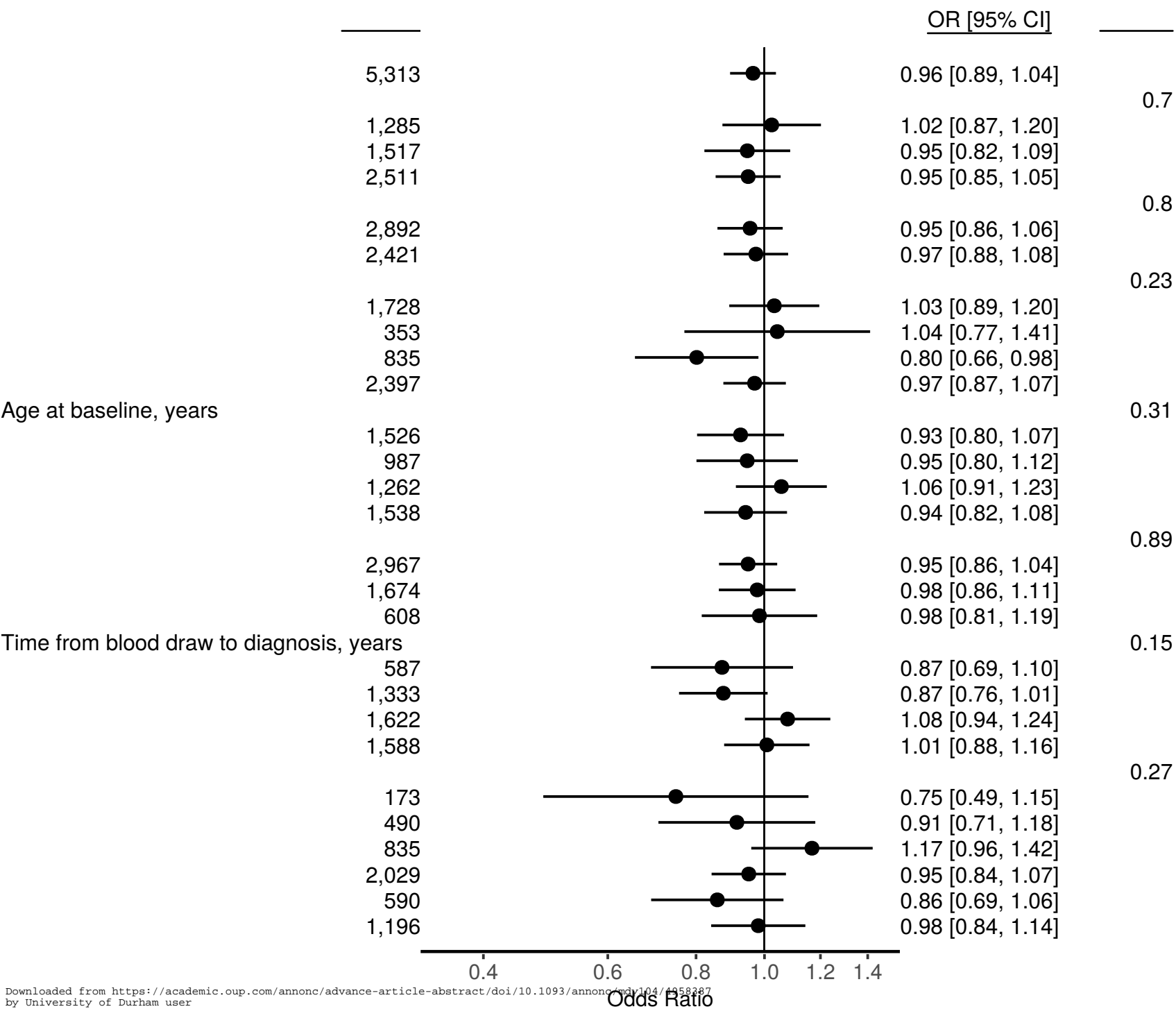


Table 1: Distribution of demographic and individual characteristics of cases and controls, overall and by continent.

| | | control | | case | | Asia | | | | Australia | | | | Europe | | | | United States | | | |
|----------------------------------|-------------|-----------|---------|-----------|---------|-----------|---------|-----------|---------|-----------|---------|-----------|---------|-----------|---------|-----------|---------|---------------|---------|------|-----|
| | | n percent | | n percent | | n percent | | n percent | | n percent | | n percent | | n percent | | n percent | | n percent | | | |
| | | n | percent | n | percent | n | percent | n | percent | n | percent | n | percent | n | percent | n | percent | n | percent | | |
| Total | | 5313 | 100 | 5313 | 100 | 1728 | 100 | 1728 | 100 | 353 | 100 | 353 | 100 | 835 | 100 | 835 | 100 | 2397 | 100 | 2397 | 100 |
| Men | | 2892 | 54 | 2892 | 54 | 1214 | 70 | 1214 | 70 | 212 | 60 | 212 | 60 | 475 | 57 | 475 | 57 | 991 | 41 | 991 | 41 |
| Women | | 2421 | 46 | 2421 | 46 | 514 | 30 | 514 | 30 | 141 | 40 | 141 | 40 | 360 | 43 | 360 | 43 | 1406 | 59 | 1406 | 59 |
| Never | | 1285 | 24 | 1285 | 24 | 560 | 32 | 560 | 32 | 49 | 14 | 49 | 14 | 107 | 13 | 107 | 13 | 569 | 24 | 569 | 24 |
| Former | | 1517 | 29 | 1517 | 29 | 176 | 10 | 176 | 10 | 145 | 41 | 145 | 41 | 190 | 23 | 190 | 23 | 1006 | 42 | 1006 | 42 |
| Current | | 2511 | 47 | 2511 | 47 | 992 | 57 | 992 | 57 | 159 | 45 | 159 | 45 | 538 | 64 | 538 | 64 | 822 | 34 | 822 | 34 |
| Age at recruitment (years) | [17,55) | 1547 | 29 | 1526 | 29 | 427 | 25 | 422 | 24 | 87 | 25 | 89 | 25 | 251 | 30 | 251 | 30 | 782 | 33 | 764 | 32 |
| | [55,60) | 995 | 19 | 987 | 19 | 368 | 21 | 355 | 21 | 57 | 16 | 64 | 18 | 198 | 24 | 191 | 23 | 372 | 16 | 377 | 16 |
| | [60,65) | 1247 | 23 | 1262 | 24 | 427 | 25 | 456 | 26 | 125 | 35 | 111 | 31 | 227 | 27 | 228 | 27 | 468 | 20 | 467 | 19 |
| | [65,86] | 1524 | 29 | 1538 | 29 | 506 | 29 | 495 | 29 | 84 | 24 | 89 | 25 | 159 | 19 | 165 | 20 | 775 | 32 | 789 | 33 |
| | no | 3963 | 75 | 3951 | 74 | 1413 | 82 | 1413 | 82 | 337 | 95 | 340 | 96 | 656 | 79 | 667 | 80 | 1557 | 65 | 1531 | 64 |
| | yes | 1350 | 25 | 1362 | 26 | 315 | 18 | 315 | 18 | 16 | 5 | 13 | 4 | 179 | 21 | 168 | 20 | 840 | 35 | 866 | 36 |
| Season-adjusted 25(OH)D (nmol/L) | [7,42) | 1328 | 25 | 1370 | 26 | 690 | 40 | 703 | 41 | 69 | 20 | 62 | 18 | 132 | 16 | 163 | 20 | 437 | 18 | 442 | 18 |
| | [42,56.6) | 1328 | 25 | 1322 | 25 | 511 | 30 | 494 | 29 | 99 | 28 | 105 | 30 | 229 | 27 | 255 | 31 | 489 | 20 | 468 | 20 |
| | [56.6,73.6) | 1328 | 25 | 1274 | 24 | 348 | 20 | 334 | 19 | 104 | 29 | 87 | 25 | 264 | 32 | 230 | 28 | 612 | 26 | 623 | 26 |
| | [73.6,237] | 1329 | 25 | 1347 | 25 | 179 | 10 | 197 | 11 | 81 | 23 | 99 | 28 | 210 | 25 | 187 | 22 | 859 | 36 | 864 | 36 |

Table 2: Characteristics of the lung cancer cases, overall and by continent.

| | | Total | | Asia | | Australia | | Europe | | United States | |
|---------------------------|-----------------|----------|---------|----------|---------|-----------|---------|----------|---------|---------------|---------|
| | | <i>n</i> | percent | <i>n</i> | percent | <i>n</i> | percent | <i>n</i> | percent | <i>n</i> | percent |
| Histology | large cell | 173 | 3 | 15 | 1 | 31 | 9 | 15 | 2 | 112 | 5 |
| | small cell | 490 | 9 | 97 | 6 | 47 | 13 | 103 | 12 | 243 | 10 |
| | squamous cell | 835 | 16 | 318 | 18 | 67 | 19 | 162 | 19 | 288 | 12 |
| | adenocarcinoma | 2029 | 38 | 589 | 34 | 152 | 43 | 260 | 31 | 1028 | 43 |
| | other | 590 | 11 | 116 | 7 | 53 | 15 | 128 | 15 | 293 | 12 |
| | unknown/missing | 1196 | 23 | 593 | 34 | 3 | 1 | 167 | 20 | 433 | 18 |
| Time to diagnosis (years) | [0,2) | 587 | 11 | 258 | 15 | 35 | 10 | 50 | 6 | 244 | 10 |
| | [2,5) | 1333 | 25 | 457 | 26 | 45 | 13 | 96 | 11 | 735 | 31 |
| | [5,10) | 1622 | 31 | 592 | 34 | 101 | 29 | 266 | 32 | 663 | 28 |
| | [10,36] | 1588 | 30 | 421 | 24 | 172 | 49 | 423 | 51 | 572 | 24 |

Table 3: Odds ratios [95% confidence intervals] for lung cancer by group of seasonally adjusted 25(OH)D, and for a doubling in 25(OH)D.

| | N controls | N cases | OR | 95% CI | <i>p</i> |
|------------------------------|------------|---------|------|--------------|----------|
| Categorical | | | | | |
| [7,25) | 245 | 283 | 1.00 | | 0.59 |
| [25,50) | 1820 | 1825 | 0.91 | [0.75, 1.1] | |
| [50,75) | 2014 | 1940 | 0.89 | [0.73, 1.08] | |
| [75,237] | 1234 | 1265 | 0.93 | [0.75, 1.15] | |
| Categorical (fourths) | | | | | |
| [7,42) | 1328 | 1370 | 1.00 | | 0.93 |
| [42,56.6) | 1328 | 1322 | 0.98 | [0.87, 1.1] | |
| [56.6,73.6) | 1328 | 1274 | 0.96 | [0.85, 1.09] | |
| [73.6,237] | 1329 | 1347 | 0.99 | [0.87, 1.13] | |
| Continuous | | | | | |
| Doubling in concentration | 5313 | 5313 | 0.98 | [0.91, 1.06] | 0.58 |

Estimates from conditional logistic regression models conditioned on matched case set, and adjusted for four groups of circulating cotinine concentration. *P*-values are from likelihood ratio tests of the 25(OH)D terms.