Circulating Folate, Vitamin B6 and Methionine in relation to Lung Cancer Risk in the Lung Cancer Cohort Consortium (LC3)

Anouar Fanidi[1]*, David Muller[1,2]*, Jian-Min Yuan[3,4]*, Victoria L. Stevens[5]*, Stephanie J. Weinstein[6]*, Demetrius Albanes[6]*, Ross Prentice[7]*, Cynthia A. Thomsen[8], Mary Pettinger[7], Qiuyin Cai[9], William J. Blot[9,10], Jie Wu[9], Alan A. Arslan[11], Anne Zeleniuch-Jacquotte[12],, Marjorie L. McCullough[5], Loic Le Marchand[13], Lynne R. Wilkens[13], Christopher A. Haiman[13], Xuehong Zhang[14], Jiali Han[15], Meir J. Stampfer[14,15,16], Stephanie A. Smith-Warner[15,16], Edward Giovannucci[14,15,16], Graham G. Giles[17,18], Allison M Hodge[17], Gianluca Severi[17,19,20,21], Mikael Johansson[22], Kjell Grankvist[22], Arnulf Langhammer[23], Steinar Krokstad[23], Marit Næss[23], Renwei Wang[3], Yu-Tang Gao[24], Lesley M. Butler[3,4], Woon-Puay Koh[25], Xiao-Ou Shu[26], Yong-Bing Xiang[27], Honglan Li[27], Wei Zheng[26], Qing Lan[27], Kala Visvanathan[28], Judith Hoffman Bolton[28], Per Magne Ueland[29,30], Øivind Midttun[31], Arve Ulvik[31], Neil E. Caporaso[6], Mark Purdue[6], Regina G. Ziegler[6], Neal D. Freedman[6], Julie E. Buring[32,33,34] , I-Min Lee[32,34], Howard D. Sesso[32,33,34], J. Michael Gaziano[32,33,35], Jonas Manjer[36], Ulrika Ericson[37], Caroline Relton[38,39] Paul Brennan[1]**, Mattias Johansson[1]**

*Co-first authors

**These authors jointly supervised this work

Corresponding author:

Paul Brennan

International Agency for Research on Cancer (IARC/WHO)

150 cours Albert Thomas

69372 Lyon cedex 08, France

Phone: +33 (0) 4 72 73 83 91

Fax: +33 (0) 4 72 73 83 42

Email: gep@iarc.fr

Affiliations:

1. Genetic Epidemiology Group, International Agency for Research on Cancer, Lyon, France
2. Department of Epidemiology and Biostatistics, Imperial College London

3. Division of Cancer Control and Population Sciences, University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania, USA.
4. Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
5. Epidemiology Research Program, American Cancer Society, Inc. 250 Williams St. Atlanta, GA 30303
6. Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH
7. Division of Public Health Sciences Fred Hutchinson Cancer Research Center 1100 Fairview Ave. N, M3-A410 Seattle, Washington 98109, U.S.A.
8. Health Promotion Sciences, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, Arizona
9. Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center and Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA
10. Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA. International Epidemiology Institute, Rockville, MD, USA
11. Departments of Obstetrics and Gynecology, Population Health and Environmental Medicine, New York University School of Medicine
12. Department of Population Health, New York University School of Medicine
13. Epidemiology program, Cancer Research Center of Hawaii, University of Hawaii
14. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA.
15. Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA.
16. Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA.
17. Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia.
18. Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Victoria, Australia
19. Molecular end Epidemiology Unit, HuGeF, Human Genetics Foundation, Torino, Italy.
20. Inserm (Institut National de la Sante et de la Recherche Medicale), Centre for Research in Epidemiology and Population Health, Villejuif 94805, France.
21. Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia
22. Umeå, University, Umeå, Sweden
23. HUNT Research Centre, Department of Public Health and General Practice, NTNU, Norwegian University of Science and Technology, Levanger, Norway
24. Department of Epidemiology, Shanghai Cancer Institute, Shanghai Jiaotong University, Shanghai, China
25. Duke-NUS Graduate Medical School Singapore, Singapore
26. Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA
27. Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA

28. Johns Hopkins University Bloomberg School of Public Health Department of Epidemiology George W Comstock Center for Public Health Research and Prevention Health Monitoring Unit, USA
29. Department of Clinical Science, University of Bergen, Bergen, Norway Laboratory of Clinical Biochemistry,
30. Haukeland University Hospital, Bergen, Norway
31. Bevital AS, 5021 Bergen, Norway
32. Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA USA.
33. Division of Aging, Department of Medicine, Brigham and Women's Hospital, Boston, MA USA.
34. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA USA
35. Boston VA Medical Center, Boston, MA USA
36. Department of Surgery, Skåne University Hospital Malmö Lund University, Malmö Sweden
37. Department of clinical sciences Malmö, Lund University, Sweden
38. Institute of Genetic Medicine, Newcastle University, Newcastle, United Kingdom
39. MRC Integrative Epidemiology Unit, School of Social & Community Medicine, University of Bristol, Bristol, United Kingdom

Keywords: one-carbon metabolism, lung cancer, consortium, prospective study

Abbreviations:

CI	confidence interval
EPIC	European Prospective Investigation into Cancer and Nutrition
OCM	one-carbon metabolism
OR	odds ratio
LC3	Lung Cancer Cohort Consortium

Article category: Research article - Epidemiology

Number of words - article: 2,998

Number of Tables: 4

Number of Figures: 1

Number of Supplementary Tables: 4

Number of Supplementary Figures: 14

Number of References: 24

Abstract

Background – Circulating concentrations of B-vitamins and factors related to one-carbon metabolism have been found to be strongly inversely associated with lung cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. The extent to which these associations are present in other study populations is unknown.

Methods – Within 20 prospective cohorts from the NCI Cohort Consortium, a nested case-control study was designed including 5,364 incident lung cancer cases and 5,364 controls that were individually matched to cases by age, sex, cohort, and smoking status. Centralized biochemical analyses were performed to measure circulating concentrations of vitamin B6, folate and methionine, as well as cotinine as an indicator of recent tobacco exposure. The association between these biomarkers and lung cancer risk was evaluated using conditional logistic regression models.

Results – Participants with higher circulating concentrations of vitamin B6 and folate had a modestly decreased risk of lung cancer risk overall, the odds ratio when comparing the top and bottom fourths [OR_{4vs1}] being 0.88 (95% confidence interval [95% CI], 0.78-1.00) and 0.86 (95% CI 0.74-0.99), respectively. We found stronger associations among for men (OR_{4vs1} 0.74; 95% CI 0.62-0.89 for vitamin B6, OR_{4vs1} 0.75; 95% CI 0.61-0.93 for folate) and ever smokers (OR_{4vs1} 0.79; 95% CI 0.68-0.91 for vitamin B6, OR_{4vs1} 0.85; 95% CI 0.72-0.99 for folate). We further noted that the association of folate was restricted to Europe/Australia and Asia, whilst no clear association was observed for the United-States. Circulating concentrations of methionine were not associated with lung cancer risk overall or in important subgroups.

Conclusion – Although confounding by tobacco exposure or reverse causation cannot be ruled out, these study results are compatible with a small decrease in lung cancer risk in ever smokers who avoid low concentrations of circulating folate and vitamin B6.

Introduction

The most important and effective means for lung cancer control is by reducing the number of people who smoke tobacco products. However, even among subjects who have quit smoking, the lifetime risk of lung cancer remains high,[1-3] and in some regions where anti-tobacco campaigns have been successful, such as the US, over 50% of lung cancer cases now occurs among former and never smokers. Given these considerations, exploring additional means of primary prevention of lung cancer is important for subgroups of high risk–in particular former smokers–who seek additional means to further reduce their risk.

One-carbon metabolism (OCM) encompasses a series of biochemical reactions involving B-vitamins that are essential to ensure balanced DNA synthesis and methylation.[4,5] Changes in specific OCM factors have been implicated in cancer development, as well as other non-communicable diseases, including neural tube defect, cardiovascular diseases, inflammation and dementia.[4,6-9] If imbalances in B-vitamins such as folate (vitamin B9) and vitamin B6 are proved to be causally implicated in lung cancer etiology, they would provide an appealing target for chemoprevention as they are modifiable by changes in diet or supplementation.[10,11]

Only two prospective studies have been published on lung cancer and circulating biomarkers of OCM, the EPIC study[12] and the Alpha-Tocopherol, Beta-Carotene, Prevention Cancer (ATBC) study of smoking men,[13] both of which reported that subjects in the upper categories of circulating vitamin B6 concentrations had approximately half the risk of lung cancer compared to those in the lower concentration categories. In particular, the EPIC nested case-control study of 899 cases and 1,770 controls reported that the inverse association of vitamin B6 with lung cancer risk was strong and consistent regardless of tobacco smoking history, and also reported strong inverse associations of methionine and folate with risk.[12]

In order to provide conclusive evidence on the importance of circulating B-vitamins and lung cancer risk, we initiated the Lung Cancer Cohort Consortium (LC3) with a combined cohort population of over 2,000,000 participants from North America, Europe, Asian, and Australia, to retrieve blood samples and conduct biochemical analysis on over 5,000 case-controls pairs.

Material and Methods

Study population

We invited all prospective cohort studies with cryopreserved baseline plasma and serum samples that in 2009 were members in the US National Cancer Institute (NCI) Cohort Consortium to participate in the study. Twenty cohorts fulfilled those criteria and accepted to participate, resulting in a combined cohort population of over 2,000,000 participants from North America, Europe, Asia and Australia. Written informed consent was provided by all study participants and the research was approved by the relevant institutional review boards. Brief details on design of the cohorts and their follow-up procedures are provided in the Supplementary methods.

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. We selected a total 5,545 lung cancer cases, and in order optimize the statistical power in smoking stratified analyses, never and former smoking cases were oversampled. For each case, one control was matched by cohort, sex, race (US cohorts only), 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), 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 years, ≥15 cigarettes per day). After various exclusions (see Supplementary methods), 5,364 lung cancer case-control pairs remained eligible for the risk analysis.

Biochemical analyses

Concentrations of vitamin B6, methionine and cotinine were determined by mass spectrometry based methods (LC-MS/MS, GC-MS/MS),[14,15] and microbiological methods were used to determine concentrations of folate (Lactobacillus casei).[16] Further details are provided in Supplementary methods.

Statistical analyses

Hierarchical linear models were used to describe the variation in average biomarker concentrations between the cohorts, and the extent to which these could be explained by differences in baseline characteristics. Relative risks of lung cancer were estimated by calculating odds ratios (OR) and 95% confidence intervals (CI) using conditional logistic regression (conditioning on matched case-control sets) with the first quartile as the referent. Additional covariates were included to account for confounding by risk factors, including indicators of educational attainment (in six categories) and tobacco exposure (in addition to matching to smoking status by design: cotinine concentrations [quartiles defined in current smokers]). Including additional covariates of body mass index (BMI) and alcohol intake did not appreciably alter the results and were not included in the final models. As a sensitivity analysis, we fitted models that were additionally adjusted for smoking duration or pack-years of smoking among ever smokers. All risk analyses were conducted overall, and stratified by smoking status and region (US, EU/AU, Asia).

As an indication of the overall statistical strength of association between each biomarker and risk, we calculated a p-value for trend by including the base-2 logarithm (log2) of the biomarker concentration as a continuous variable in a separate conditional logistic regression model. The same approach was used in stratified risk analyses according to other pre-defined demographic characteristics and risk factors. OR estimates per log2 unit (log_2 OR) may be interpreted as the relative risk associated with a doubling in the concentration of a circulating biomarker.

All statistical analyses were conducted using SAS 9.2 (Cary, North Carolina),[17] R version 3.1.3[18] or Stan version 2.9.0.[19]

Additional details on the statistical methods are provided in the Supplementary methods.

Results

Baseline Characteristics

Baseline characteristics of the study subjects separated by geographic areas are shown in Table 1. The final study population included 5,364 lung cancer cases and 5,364 individually

matched controls. Median age at cohort enrolment was 63 years and the median time between blood draw and lung cancer diagnosis was 6.3 years. Because we oversampled never and former smokers, approximately half of the study population were current smokers at recruitment (2,519 case-control pairs, 47%), whereas the other half were either former (1,518 case-control pairs, 28%) or never smokers (1,327 case-control pairs, 25%). Overall, 46% of the participants were women, though the sex distribution varied by region, the US cohorts contributing more women and the cohorts from Europe and Asia contributed more men.

Variation in circulating biomarkers

There was substantial between-cohort variation in concentrations of all biomarkers (Table 1), in particular for folate and vitamin B6 concentrations that were substantially higher in US cohorts than for EU/AU or Asian cohorts. This could only partly be explained by differences in age, sex, smoking status, or BMI (Figure 1). For circulating vitamin B6 and methionine, the between-cohort variability was similar before and after covariate adjustment (Figure 1). Conversely, for circulating folate, adjustment for age, sex, smoking status, and BMI reduced the between-cohort variability, the sd being 0.70 (90% CI 0.54-0.97) without adjustment and 0.44 (90% CI 0.32-0.65) with adjustment. Limited data on supplementation use suggested that a part of the differences in vitamin B6 and folate concentrations between US and non-US cohorts can be explained by multivitamin use (vitamin B6 median concentrations among US multivitamin users: 76.2 nmol/L, non-users: 39.2 nmol/L; folate median concentrations users:46.5 nmol/L, non-users:25.5 nmol/L, Supplementary Table 5).

Risk analysis overall and stratified by region

Participants with higher circulating vitamin B6 and folate had an approximate 20% lower risk of lung cancer overall (Table 2 and 3), the odds ratio when comparing the top and bottom quartiles [OR_{4vs1}] being 0.81 for vitamin B6 (95% CI 0.72-0.92) and 0.80 for folate (95% CI 0.70-0.92). These OR estimates were slightly attenuated after adjustment for circulating cotinine and education (OR_{4vs1} 0.88, 95% CI 0.78-1.00 for vitamin B6 and OR_{4vs1} 0.86, 95% CI 0.74-0.99 for folate). We did not observe any clear association between circulating methionine and risk of lung cancer (OR_{4vs1} 0.95, 95% 0.85-1.07, Table 4), apart from among US men (OR_{4vs1} 0.68, 95% 0.52-0.89, Table 4). Among participants with available information

on smoking duration and pack-years of smoking (90% of ever smokers), further adjustment by continuous number of years of smoking or pack-years of smoking did not notably affect the estimates (Supplementary Table 1-4).

In risk analyses stratified by region (Table 2-4), the overall inverse associations of vitamin B6 and folate were mainly driven by Asian and EU/AU participants. The adjusted OR_{4vs1} for vitamin B6 was 0.82 in the Asian cohorts (95% CI 0.65-1.02) and 0.78 for the EU/AU cohorts (95% CI 0.62-1.00), and 0.95 for the US cohorts (95% CI 0.80-1.13). Similarly, for folate, the adjusted OR_{4vs1} was 0.84 for the Asian cohorts (95% CI 0.68-1.03) and 0.77 for the EU/AU cohorts (95% CI 0.60-0.99), and 0.95 for the US cohorts (95% CI 0.78-1.15).

Risk analysis stratified by smoking status and sex

In analyses stratified by smoking status, we observed an inverse association between vitamin B6 and overall risk for former (OR_{4vs1} 0.72, 95% CI 0.57-0.91) and current smokers (OR_{4vs1} 0.79, 95% CI 0.65-0.95) (Table 2). Folate was also inversely associated with risk for former (OR_{4vs1} 0.66, 95% CI 0.51-0.85), but not for current smokers (OR_{4vs1} 0.97, 95% CI 0.77-1.21) (Table 3). When analyzing ever smokers overall (i.e. former and current smokers combined), we observed an inverse association for circulating vitamin B6 (OR_{4vs1} 0.79, 95% CI 0.88-0.91) and folate (OR_{4vs1} 0.85, 95% CI 0.72-0.99). Of note, whilst no clear association was seen between folate and risk for never smokers (OR_{4vs1} 0.86, 95% CI 0.63-1.17), we observed a positive association between vitamin B6 and risk for never smokers, the OR_{4vs1} being 1.51 (95% CI 1.14-2.01) (Table 2).

Further analyses stratified by sex showed that any inverse association between folate or vitamin B6 and lung cancer risk was restricted to ever smokers and men. First, the association of vitamin B6 in men (OR_{4vs1} 0.74, 95% CI 0.62-0.89) was consistently observed in cohorts from each region (Table 2), with similar observations for folate (overall OR_{4vs1} 0.75; 95% CI 0.61-0.93) (Table 3). Conversely, no association was seen among for women overall, nor for women separated by region (Table 2-3).

No clear differences in OR estimates were seen in risk analyses conducted on cohorts providing plasma compared to cohorts providing serum ($P_{heterogeneity}$ >0.33).

Stratified risk analysis by demographic and diagnostic parameters

To assess potential of effect modification, we fitted models that included interaction terms between circulating folate or vitamin B6 and various covariates (Supplementary Figures 1-8). Gender modified the associations between circulating concentrations of vitamin B6 and folate and lung cancer risk ($P_{heterogeneity}$ ≤0.05, Supplementary Tables 1 and 4). When assessing blood samples taken up to 10 years prior to diagnosis, we observed that the overall inverse association between vitamin B6 and lung cancer risk was driven by cases diagnosed closer to blood draw ($P_{heterogeneity}$ = 0.001, Supplementary Figure 1), as well as by former (OR_{log2} 0.81; 95% CI 0.70-0.90, Supplementary Figure 3) and current smokers (OR_{log2} 0.79; 95% CI 0.69-0.91, Supplementary Figure 4). Stratified analyses by histology showed a stronger inverse association between vitamin B6 and folate and risk of lung squamous cell cancer (OR_{log2} 0.86; 95% CI 0.77-0.95 for vitamin B6 and OR_{log2} 0.88; 95% CI 0.78-0.98 for folate) than for other histological types (Supplementary Figures 1 and 5). Stratified analysis by period of blood collection before and after 1996 when folate fortification of food items took effect in US did not reveal any discernable difference in lung cancer risk for circulating folate (Supplementary Figure 12), nor did analysis stratified on multivitamin use (data not shown). The associations between biomarker concentrations and lung cancer risk for individual cohorts are presented in Supplementary Figures 9-11. We also estimated the potential risk increase associated with being clinically deficient in vitamin B6 and folate, and observed a 23% increase in risk of lung cancer for those deficient in vitamin B6 (OR = 1.23, 95% CI 1.09-1.38) that was more apparent for being diagnosed with squamous cell cancer (OR=1.54, 95% CI 1.18-2.02), and also those for developing lung cancer within the first 3 years following blood collection (OR=1.83, 95% CI 1.42-2.36) (Supplementary Figure 13). Similarly, for folate we observed an increased risk for those classified as deficient (OR=1.27, 95% CI 1.07-1.49, Supplementary Figure 14). Further stratified analyses are presented in Supplementary Figures 13 and 14.

Discussion

Our aim was to evaluate the associations between lung cancer risk and circulating concentrations of folate, vitamin B6, and methionine, within a consortium of 20 prospective cohorts from the US, Europe, Asia and Australia. We found that individuals with low

concentrations of vitamin B6 and folate have a small risk increase of lung cancer, particularly among men, current and former smokers, and participants living outside the United States.

Previous studies and the Lung Cancer Cohort Consortium (LC3)

To date, this is the largest cohort consortium initiative for which circulating biomarkers where assessed in relation to a single cancer outcome with coordinated biochemical analysis conducted in one centralized laboratory. The background to this study was the initial findings from the EPIC cohort involving 899 lung cancer cases and 1,770 controls.[12] The EPIC study reported a strong inverse association between vitamin B6 and lung cancer regardless of smoking history, and additionally, found similar inverse associations with risk for folate and methionine. The results of the EPIC study were noteworthy as they suggested a potentially important role of the one-carbon metabolism pathway in lung cancer etiology, in addition to tobacco exposure. Along with additional evidence from the ATBC cohort,[13] this motivated the organization of the Lung Cancer Cohort Consortium (LC3) in order to conclusively evaluate the extent to which these findings translate into other populations from Asia, the United States, Europe and Australia.

Interpreting the results

The current study discerned relatively weak inverse associations between folate, vitamin B6 and lung cancer risk overall, associations that differed substantially between groups of different tobacco exposure history. For former and current smokers, we observed slightly lower risk among for those study participants with higher concentrations of folate and vitamin B6 compared to those having lower concentrations. This inverse association for lung cancer risk with vitamin B6 or folate can also be interpreted as an increased risk among participants in the lowest category of concentrations compared to those in the remaining three higher categories (Tables 2 and 3). In contrast, methionine was inversely associated with risk in US men, whereas no associations were noted overall or in other important subgroup.

Previous studies have shown that being a current smoker is clearly associated with lower circulating concentrations of folate and vitamin B6,[12] and our analysis confirms this relation

(data not shown). This highlights the importance of carefully accounting for tobacco exposure in risk analyses of B-vitamins and lung cancer. We used circulating cotinine, an objective and accurate measure of current tobacco exposure,[20] and the weak inverse associations of vitamin B6 and folate with lung cancer risk remained after adjusting for cotinine. Further accounting for smoking duration and intensity did not substantively alter the OR estimates. In contrast to current smokers, former smokers tend to have similar circulating B-vitamin concentrations as never smokers.[12] This suggests that the potential confounding effect of tobacco exposure on the associations between circulating vitamin B6/folate and lung cancer risk is of less concern for former smokers than for current smokers. Despite these considerations, and given the important impact of both past and current tobacco exposure on lung cancer risk, our results do not allow ruling out confounding by tobacco smoking as a possible explanation for the observed associations.

Another potential explanation for the observed association of vitamin B6 with risk is reverse causation. In analysis stratified on smoking status, we observed a gradually stronger inverse association between vitamin B6 and risk of lung cancer in both former and current smokers when measured in blood drawn closer to diagnosis (Supplementary Figure 1, 3 and 4), results that would be consistent with pre-clinical metabolic changes due to the underlying disease progression. In contrast, no such clear relation by time from blood draw to diagnosis was observed for folate.

In interpreting these results, we note that they are compatible with a small benefit in terms of lung cancer risk for smokers who avoid low circulating concentrations of vitamin B6 and folate. However, we cannot rule out the possibility that residual confounding by tobacco exposure and/or reverse causation due to the disease progression, underlie their associations with risk.

The most notable difference between our results and those from the previous EPIC study was for never smokers, where vitamin B6 was positively associated with risk among for women in the current study, the opposite of what was observed in the previous EPIC study. The reason for this stark discrepancy is unclear and not explained by differences in study characteristics.

Study strengths and limitations

Strengths of this study included the large study sample of 5,364 case-control pairs, the prospective study design and the use of pre-diagnostic plasma or serum, the centralized biochemical analysis with a robust quality control protocol, and the wide participation of cohorts from different geographical regions. We benefitted from over-sampling never and former smokers, as well as matching by history of tobacco exposure, thus allowing for well-powered stratified analysis. However, this oversampling also meant that the individual cohorts provided different proportions of subjects by smoking status. For instance, the WHI only contributed never smoking women and the ATBC study only contributed current smoking men. Another particular feature of this study was that circulating vitamin B6 and folate varied substantially across cohorts and continents. Indeed, US participants had about 65% higher B6 and almost three-fold the median folate concentrations of European and Asian study participants (see Table 1). A likely explanation of this is that up to 50% of U.S. citizens aged 50 years or more regularly consume dietary supplements such as multi-vitamin pills,[21] and because of the folate fortification in the US.[22] Differences in baseline characteristics only partially accounted for these differences, and this means that we could not efficiently compare participants at extremes of the distribution of folate concentrations because the controls were matched to cases within the same cohort, and we were reluctant to break that matching.

Study implications and conclusions

This study highlights the importance of re-evaluating promising associations of putative risk factors indicated in initial studies across multiple study populations in a coordinated fashion. The LC3 clearly demonstrates that it is feasible to gather cohorts from around the world and provide robust information on disease risk association before resorting to intervention studies that are costly and risk harming the study participants. Whilst our results were compatible with a modest decrease in lung cancer risk among former and current smokers who avoid low circulating concentrations of vitamin B6 and folate, we could not rule out the possibility that confounding due to tobacco or reverse causation may have given rise to spurious associations. Given that any potential beneficial effect of vitamin B6 or folate–if real–is likely to be small, our findings do not support the

conduct of additional prevention studies with the view of using B-vitamins for primary prevention of lung cancer.

References

1. Brennan P, Crispo A, Zaridze D, et al. High cumulative risk of lung cancer death among smokers and nonsmokers in Central and Eastern Europe. *Am J Epidemiol.* 2006;164(12):1233-1241.
2. Crispo A, Brennan P, Jockel KH, et al. The cumulative risk of lung cancer among current, ex- and never-smokers in European men. *Br J Cancer.* 2004;91(7):1280-1286.
3. Ebbert JO, Yang P, Vachon CM, et al. Lung cancer risk reduction after smoking cessation: observations from a prospective cohort of women. *J Clin Oncol.* 2003;21(5):921-926.
4. Kim YI. Folate and colorectal cancer: an evidence-based critical review. *Molecular nutrition & food research.* 2007;51(3):267-292.
5. Locasale JW. Serine, glycine and one-carbon units: cancer metabolism in full circle. *Nature reviews Cancer.* 2013;13(8):572-583.
6. Kim YI. Nutritional epigenetics: impact of folate deficiency on DNA methylation and colon cancer susceptibility. *The Journal of nutrition.* 2005;135(11):2703-2709.
7. Vollset SE, Clarke R, Lewington S, et al. Effects of folic acid supplementation on overall and site-specific cancer incidence during the randomised trials: meta-analyses of data on 50,000 individuals. *Lancet.* 2013;381(9871):1029-1036.
8. Ueland PM, McCann A, Midttun O, Ulvik A. Inflammation, vitamin B6 and related pathways. *Mol Aspects Med.* 2016.
9. Bailey LB, Stover PJ, McNulty H, et al. Biomarkers of Nutrition for Development-Folate Review. *The Journal of nutrition.* 2015;145(7):1636S-1680S.
10. Galluzzi L, Vacchelli E, Michels J, et al. Effects of vitamin B6 metabolism on oncogenesis, tumor progression and therapeutic responses. *Oncogene.* 2013;32(42):4995-5004.
11. Kalmbach RD, Choumenkovitch SF, Troen AM, D'Agostino R, Jacques PF, Selhub J. Circulating folic acid in plasma: relation to folic acid fortification. *Am J Clin Nutr.* 2008;88(3):763-768.
12. Johansson M, Relton C, Ueland PM, et al. Serum B vitamin levels and risk of lung cancer. *JAMA.* 2010;303(23):2377-2385.
13. Hartman TJ, Woodson K, Stolzenberg-Solomon R, et al. Association of the B-vitamins pyridoxal 5'-phosphate (B(6)), B(12), and folate with lung cancer risk in older men. *American journal of epidemiology.* 2001;153(7):688-694.
14. Midttun O, Hustad S, Ueland PM. Quantitative profiling of biomarkers related to B-vitamin status, tryptophan metabolism and inflammation in human plasma by liquid chromatography/tandem mass spectrometry. *Rapid communications in mass spectrometry : RCM.* 2009;23(9):1371-1379.
15. Ueland PM, Midttun O, Windelberg A, Svardal A, Skalevik R, Hustad S. Quantitative profiling of folate and one-carbon metabolism in large-scale epidemiological studies by mass spectrometry. *Clinical chemistry and laboratory medicine : CCLM / FESCC.* 2007;45(12):1737-1745.
16. Molloy AM, Scott JM. Microbiological assay for serum, plasma, and red cell folate using cryopreserved, microtiter plate method. *Methods in enzymology.* 1997;281:43-53.
17. *SAS® 9.4 Help and documentation* [computer program]. Version 9.4. Cary, NC, USA: SAS Institute Inc; 2014.
18. *RStan: the R interface to Stan* [computer program]. Version Version 2.52014.
19. *Stan Modeling Language Users Guide and Reference Manual* [computer program]. Version Version 2.5.0.2014.
20. Boffetta P, Clark S, Shen M, Gislefoss R, Peto R, Andersen A. Serum cotinine level as predictor of lung cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* 2006;15(6):1184-1188.

21. Bailey RL, Gahche JJ, Lentino CV, et al. Dietary supplement use in the United States, 2003-2006. *The Journal of nutrition.* 2011;141(2):261-266.
22. Jacques PF, Selhub J, Bostom AG, Wilson PW, Rosenberg IH. The effect of folic acid fortification on plasma folate and total homocysteine concentrations. *N Engl J Med.* 1999;340(19):1449-1454.

Funding

The Lung Cancer Cohort Consortium (LC3) was supported by Grant NIH / NCI (n° 1U01CA155340-01). The funding organization 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.

Drs Ueland and Midttun 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.

PB and MJ initiated, acquired the main funding, and designed this investigation. PMU and ØM led the laboratory analysis. AF and DM conducted the statistical analysis under supervision of MJ and PB. AF, MJ, JMY, VLS, SJW, DA, DM, RP, and PB drafted the first version of the manuscript. All authors were involved with collection of data, data interpretation, critical revisions of the paper, and approval of the final version. PB had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. PB and MJ are the overall coordinators of the LC3 consortium that was implemented in collaboration with the main investigators in the collaborating cohorts.

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

Figure 1. Estimated between-cohort variation in mean concentration of vitamin B6, folate, and methionine. Before (left) and after (right) adjustment for age, sex, smoking status, and BMI. Histograms show the posterior distribution of the overall mean and standard deviation (SD) of the cohort-specific intercepts. The "caterpillar plots" show the estimated cohort-specific mean concentrations, along with 50% and 90% credible intervals (thick and thin bars, respectively).

Table 1. Baseline and clinical characteristics of study participants overall and per continent

Discrete variables	US cohorts* No.(%) of participants in group		EU/AU cohorts* No.(%) of participants in group		Asian cohorts* No.(%) of participants in group		Overall No.(%) of participants in group	
	Cases (n=2400)	Matched controls (n=2400)	Cases (n=1189)	Matched controls (n=1189)	Cases (n=1775)	Matched controls (n=1775)	Cases (n=5364)	Matched controls (n=5364)
Sex								
Men	991 (41.3%)	991 (41.3%)	688 (57.9%)	688 (57.9%)	1229 (69.2%)	1229 (69.2%)	2908 (54.2%)	2908 (54.2%)
Women	1409 (58.7%)	1409 (58.7%)	501 (42.1%)	501 (42.1%)	546 (30.8%)	546 (30.8%)	2456 (45.8%)	2456 (45.8%)
Smoking status								
Never	569 (23.7%)	569 (23.7%)	156 (13.1%)	156 (13.1%)	602 (33.9%)	602 (33.9%)	1327 (24.7%)	1327 (24.7%)
Former	1007 (42%)	1007 (42%)	335 (28.2%)	335 (28.2%)	176 (9.9%)	176 (9.9%)	1518 (28.3%)	1518 (28.3%)
Current	824 (34.3%)	824 (34.3%)	698 (58.7%)	698 (58.7%)	997 (56.2%)	997 (56.2%)	2519 (47%)	2519 (47%)
Education								
less than high school	237 (9.9%)	215 (9%)	661 (55.6%)	596 (50.2%)	898 (50.6%)	883 (49.7%)	1797 (33.5%)	1696 (31.6%)
completed high school	357 (14.9%)	374 (15.6%)	159 (13.4%)	180 (15.2%)	243 (13.7%)	230 (13%)	759 (14.1%)	784 (14.6%)
vocational school	422 (17.6%)	435 (18.1%)	180 (15.2%)	200 (16.8%)	289 (16.3%)	279 (15.7%)	891 (16.6%)	914 (17%)
some college	402 (16.8%)	393 (16.4%)	107 (9%)	129 (10.9%)	171 (9.6%)	196 (11%)	680 (12.7%)	718 (13.4%)
college graduate	357 (14.9%)	319 (13.3%)	63 (5.3%)	64 (5.4%)	104 (5.9%)	113 (6.4%)	524 (9.8%)	496 (9.2%)
graduate studies	574 (23.9%)	637 (26.5%)	10 (0.8%)	8 (0.7%)	62 (3.5%)	65 (3.7%)	646 (12%)	710 (13.2%)
Unknown	51 (2.1%)	27 (1.1%)	8 (0.7%)	10 (0.8%)	8 (0.5%)	9 (0.5%)	67 (1.2%)	46 (0.9%)
Alcohol use (any type)								
Never	250 (10.4%)	252 (10.5%)	292 (24.6%)	286 (24.1%)	946 (53.3%)	974 (54.9%)	1488 (27.7%)	1512 (28.2%)
Ever	2013 (83.9%)	2041 (85%)	643 (54.1%)	643 (54.1%)	829 (46.7%)	801 (45.1%)	3485 (65%)	3485 (65%)
Unknown	137 (5.7%)	107 (4.5%)	254 (21.4%)	260 (21.9%)	0 (0%)	0 (0%)	391 (7.3%)	367 (6.8%)
Body Mass index [kg/m^2]								
<18.5	30 (1.3%)	31 (1.3%)	14 (1.2%)	9 (0.8%)	157 (8.8%)	113 (6.4%)	201 (3.7%)	153 (2.9%)
18.5-25	1088 (45.3%)	1020 (42.5%)	521 (43.8%)	435 (36.6%)	1203 (67.8%)	1192 (67.2%)	2812 (52.4%)	2647 (49.3%)
25-30	841 (35%)	858 (35.8%)	468 (39.4%)	536 (45.1%)	369 (20.8%)	424 (23.9%)	1678 (31.3%)	1818 (33.9%)
≥30	378 (15.8%)	430 (17.9%)	185 (15.6%)	207 (17.4%)	46 (2.6%)	46 (2.6%)	609 (11.4%)	683 (12.7%)
Unknown	63 (2.6%)	61 (2.5%)	1 (0.1%)	2 (0.2%)	0 (0%)	0 (0%)	64 (1.2%)	63 (1.2%)
Continuous variables, median (5th-95th percentile)								
Age at recruitment (years)	60 (42-74)	60 (42-74)	60 (45-70)	60 (45-70)	60 (46-72)	60 (46-72)	60 (44-72)	60 (44-72)
Blood concentrations for biomakers								
Vitamin B6 (Pyridoxal 5'-phosphate), nmol/L	47.6 (15.2-266)	49.9 (16.4-271)	28.9 (12.6-127)	30.9 (13.3-102)	29.0 (11.1-114)	31.2 (12.4-119)	35.2 (12.6-205)	37.1 (13.9-197)
Vitamin B9 (folate), nmol/L	32.9 (8.3-114.2)	33.7 (8.1-114.9)	10.2 (4.4-30.1)	10.8 (4.9-32.9)	13.9 (6.2-40.5)	15.1 (6.7-40.7)	17.0 (6.0-91.8)	17.9 (6.3-90.3)
Methionine, µmol/L	25.8 (18.4-39.6)	26.2 (18.8-39.7)	26.1 (19.1-37.8)	26.3 (18.7-37.8)	27.3 (18.0-42.4)	27.4 (18.2-41.9)	26.3 (18.4-40.4)	26.6 (18.6-40.3)
Clinical characteristics, case participants only								
Age at diagnosis, median (range), years	70 (55-83)		69 (54-79)		69 (52-80)		69.8 (53.6-82.0)	
Time from blood draw to diagnosis (years)	5.2 (1-15.5)		10.0 (1.5-16.0)		5.8 (0.7-16.5)		6.3 (1.0-16.0)	
Hystology, No. (%)								
Large cell carcinoma	112 (4.6%)		46 (4%)		16 (1%)		174 (3.3%)	
Small cell carcinoma	245 (10.4%)		150 (12.5%)		99 (5.5%)		492 (9.2%)	
Squamous cell carcinoma	291 (11.9%)		231 (19.5%)		319 (17.9%)		836 (15.5%)	
Adenocarcinoma	1034 (42.7%)		419 (34.5%)		615 (34.6%)		2056 (38.4%)	
Other / Unknown	735 (31.4%)		357 (29.5%)		726 (41%)		1806 (33.6%)	

* US cohorts included 83% white, 11% African American, and 6% participants indicated as 'Other'.

Table 2. Odds ratios of lung cancer for circulating concentrations of vitamin B6

			Odds Ratio (95% Confidence Interval)						
			Cases compared to matched controls, unadjusted[a]	**Cases compared to matched controls, adjusted for cotinine and education[b]**	**Never Smokers[c]**	**Former Smokers[c]**	**Current smokers[b]**	**Men[b]**	**Women[b]**
LC3 participants	**Quartile (Range)**	**Case/Controls participants**	**(n=5364/5364)[a]**	**(n=5364/5364)[b]**	**(n=1327/1327)[c]**	**(n = 1518/1518)[c]**	**(n = 2519/2519)[b]**	**(n=2908/2908)[b]**	**(n=2456/2456)[b]**
	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]								
United States (n=11 cohorts)	1 (5.08-29.89)	629/600	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (29.90-49.90)	618/600	0.98 (0.83 - 1.15)	1.00 (0.84 - 1.18)	1.45 (0.99 - 2.13)	0.88 (0.68 - 1.14)	0.95 (0.73 - 1.24)	0.82 (0.63 - 1.06)	1.14 (0.91 - 1.41)
	3 (49.91-90.05)	574/600	0.90 (0.77 - 1.07)	0.93 (0.78 - 1.10)	1.19 (0.82 - 1.72)	0.87 (0.67 - 1.13)	0.89 (0.67 - 1.17)	0.77 (0.59 - 1.00)	1.04 (0.84 - 1.30)
	4 > 90.05	579/600	0.91 (0.76 - 1.07)	0.95 (0.8 - 1.13)	1.19 (0.82 - 1.72)	0.90 (0.69 - 1.17)	0.85 (0.63 - 1.15)	0.75 (0.56 - 0.99)	1.11 (0.89 - 1.38)
	P for trend[h]		0.14	0.40	0.77	0.15	0.52	0.02	0.47
	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][e]								
Europe /Australia (n=5 cohorts)	1 (4.37-22.41)	378/297	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.42-30.90)	269/297	0.70 (0.56 - 0.88)	0.73 (0.58 - 0.92)	1.32 (0.59 - 2.98)	0.97 (0.59 - 1.60)	0.62 (0.47 - 0.82)	0.73 (0.54 - 0.99)	0.71 (0.49 - 1.03)
	3 (30.91-44.95)	258/298	0.65 (0.51 - 0.83)	0.69 (0.54 - 0.88)	1.19 (0.56 - 2.54)	0.61 (0.37 - 0.99)	0.71 (0.52 - 0.97)	0.74 (0.53 - 1.02)	0.61 (0.42 - 0.89)
	4 > 44.96	284/297	0.73 (0.57 - 0.92)	0.78 (0.62 - 1.00)	1.57 (0.77 - 3.19)	0.73 (0.46 - 1.17)	0.76 (0.55 - 1.04)	0.68 (0.50 - 0.94)	0.91 (0.64 - 1.31)
	P for trend[h]		0.12	0.36	0.56	0.42	0.41	0.17	0.97
	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][f]								
Asia (n=4 cohorts)	1 (4.81-20.49)	528/443	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (20.50-30.82)	440/444	0.78 (0.64 - 0.96)	0.90 (0.73 - 1.10)	1.33 (0.80 - 2.21)	0.67 (0.32 - 1.41)	0.92 (0.72 - 1.17)	0.89 (0.71 - 1.11)	1.29 (0.74 - 2.25)
	3 (31.23-46.99)	398/444	0.67 (0.55 - 0.84)	0.79 (0.63 - 0.98)	1.59 (0.95 - 2.68)	0.60 (0.31 - 1.17)	0.66 (0.50 - 0.86)	0.68 (0.52 - 0.87)	1.45 (0.84 - 2.51)
	4 > 47.00	409/444	0.68 (0.55 - 0.87)	0.82 (0.65 - 1.02)	1.55 (0.99 - 2.72)	0.54 (0.28 - 1.02)	0.66 (0.48 - 0.93)	0.75 (0.58 - 0.98)	1.34 (0.79- 2.29)
	P for trend[h]		0.002	0.08	0.06	0.29	0.003	0.008	0.50
	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][g]								
All cohorts (n=20 cohorts)	1 (4.37-23.93)	1501/1341	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.94-37.05)	1319/1341	0.85 (0.76 - 0.95)	0.90 (0.81 - 1.01)	1.53 (1.15 - 2.05)	0.73 (0.58 - 0.93)	0.86 (0.74 - 1.00)	0.77 (0.66 - 0.89)	1.14 (0.96 - 1.37)
	3 (37.06-62.50)	1254/1341	0.80 (0.72 - 0.90)	0.87 (0.77 - 0.97)	1.52 (1.15 - 2.00)	0.66 (0.53 - 0.84)	0.83 (0.70 - 0.98)	0.77 (0.66 - 0.91)	1.03 (0.86 - 1.23)
	4 > 62.51	1290/1341	0.81 (0.72 - 0.92)	0.88 (0.78 - 1.00)	1.51 (1.14 - 2.01)	0.72 (0.57 - 0.91)	0.79 (0.65 - 0.95)	0.74 (0.62 - 0.89)	1.10 (0.92 - 1.31)
	P for trend[h]		$9x10^{-4}$	0.06	0.12	0.05	0.005	$2x10^{-4}$	0.37

a) Assessed by analysing lung cancer cases and their individually matched controls by conditional logistic regression, conditioning on individual case set.

b) Assessed by analysing lung cancer cases and their individually matched controls by conditional logistic regression, conditioning on individual case set, and adjusting for circulating cotinine (in quartiles) and education (in 7 categories) .

c) Assessed by analysing lung cancer cases and their individually matched controls by conditional logistic regression, conditioning on individual case set, and adjusting for education (in 7 categories) .

d) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 2400 individually matched controls (US cohorts; n=11).

e) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 1189 individually matched controls (European/Australian cohorts; n=5).

f) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 1775 individually matched controls (Asian cohorts; n=4).

g) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 5364 individually matched controls (pooled cohorts; n=20).

h) P for trend assessed by the base 2 logarithm of the circulating concentrations.

Table 3. Odds ratios of lung cancer for circulating concentrations of folate

			Odds Ratio (95% Confidence Interval)						
			Cases compared to matched controls, unadjusted[a]	**Cases compared to matched controls, adjusted for cotinine and education[b]**	**Never Smokers[c]**	**Former Smokers[c]**	**Current smokers[b]**	**Men[b]**	**Women[b]**
LC3 participants	**Quartile (Range)**	**Case/Controls participants**	**(n=5364/5364)[a]**	**(n=5364/5364)[b]**	**(n=1327/1327)[c]**	**(n = 1518/1518)[c]**	**(n = 2519/2519)[b]**	**(n=2908/2908)[b]**	**(n=2456/2456)[b]**
United States (n=11 cohorts)	**Folate (Vitamin B9) [nmol/L][d]**								
	1 (1.70-17.45)	633/600	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (17.46-33.35)	578/600	0.90 (0.77 - 1.07)	0.93 (0.78 - 1.10)	0.83 (0.54 - 1.29)	0.95 (0.72 - 1.24)	0.92 (0.72 - 1.18)	0.86 (0.66 - 1.12)	0.98 (0.79 - 1.22)
	3 (33.36-59.10)	598/600	0.93 (0.78 - 1.10)	0.95 (0.80 - 1.14)	0.87 (0.56 - 1.36)	0.86 (0.66 - 1.13)	1.05 (0.78 - 1.42)	0.92 (0.69 - 1.22)	0.98 (0.78 - 1.24)
	4 > 59.10	591/600	0.91 (0.75 - 1.10)	0.95 (0.78 - 1.15)	0.92 (0.57 - 1.47)	0.87 (0.65 - 1.16)	1.02 (0.72 - 1.45)	0.78 (0.56 - 1.07)	1.07 (0.84 - 1.37)
	P for trend[h]		0.27	0.52	0.84	0.27	0.88	0.44	0.85
Europe /Australia (n=5 cohorts)	**Folate (Vitamin B9) [nmol/L][e]**								
	1 (0.24-7.78)	348/297	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (7.79-10.77)	291/297	0.82 (0.65 - 1.03)	0.84 (0.66 - 1.06)	1.30 (0.65 - 2.60)	0.71 (0.44 - 1.14)	0.81 (0.60 - 1.10)	0.82 (0.59 - 1.12)	0.85 (0.59 - 1.22)
	3 (10.78-15.98)	282/298	0.78 (0.61 - 0.98)	0.80 (0.63 - 1.02)	0.65 (0.33 - 1.25)	0.69 (0.44 - 1.10)	0.90 (0.65 - 1.23)	0.76 (0.55 - 1.04)	0.85 (0.58 - 1.23)
	4 > 15.98	268/297	0.73 (0.57 - 0.93)	0.77 (0.60 - 0.99)	0.98 (0.49 - 1.97)	0.58 (0.36 - 0.93)	0.83 (0.60 - 1.16)	0.69 (0.49 - 0.98)	0.87 (0.60 - 1.27)
	P for trend[h]		0.007	0.02	0.89	0.007	0.37	0.02	0.47
Asia (n=4 cohorts)	**Folate (Vitamin B9) [nmol/L][f]**								
	1 (0.17-10.59)	525/443	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.60-15.00)	461/444	0.84 (0.70 - 1.02)	0.88 (0.72 - 1.06)	0.96 (0.65 - 1.40)	0.40 (0.21 - 0.74)	0.96 (0.75 - 1.23)	0.85 (0.68 - 1.06)	1.01 (0.68 - 1.5)
	3 (15.01-21.00)	366/444	0.67 (0.55 - 0.81)	0.73 (0.59 - 0.89)	0.76 (0.52 - 1.10)	0.45 (0.23 - 0.85)	0.80 (0.65 - 1.12)	0.75 (0.59 - 0.95)	0.78 (0.53 - 1.14)
	4 >21.01	423/444	0.77 (0.63 - 0.94)	0.84 (0.68 - 1.03)	0.89 (0.61 - 1.29)	0.65 (0.35 - 1.18)	0.84 (0.64 - 1.11)	0.74 (0.58 - 0.94)	1.12 (0.76 - 1.65)
	P for trend[h]		0.001	0.02	0.26	0.26	0.09	0.008	0.97
All cohorts (n=20 cohorts)	**Folate (Vitamin B9) [nmol/L][g]**								
	1 (0.17-10.92)	1489/1341	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.93-17.89)	1331/1341	0.87 (0.78 - 0.97)	0.90 (0.81 - 1.01)	0.87 (0.67 - 1.13)	0.74 (0.58 - 0.95)	0.96 (0.83 - 1.12)	0.90 (0.77 - 1.04)	0.92 (0.77 - 1.10)
	3 (17.90-34.92)	1243/1341	0.79 (0.70 - 0.89)	0.84 (0.74 - 0.95)	0.81 (0.62 - 1.06)	0.73 (0.57 - 0.93)	0.88 (0.74 - 1.04)	0.77 (0.65 - 0.91)	0.94 (0.78 - 1.13)
	4 > 34.97	1301/1341	0.80 (0.70 - 0.92)	0.86 (0.74 - 0.99)	0.86 (0.63 - 1.17)	0.66 (0.51 - 0.85)	0.97 (0.77 - 1.21)	0.75 (0.61 - 0.93)	0.97 (0.79 - 1.19)
	P for trend[h]		$2x10^{-4}$	0.01	0.37	0.01	0.15	0.001	0.66

a) Assessed by analysing lung cancer cases and their individually matched controls by conditional logistic regression, conditioning on individual case set.

b) Assessed by analysing lung cancer cases and their individually matched controls by conditional logistic regression, conditioning on individual case set, and adjusting for circulating cotinine (in quartiles) and education (in 7 categories) .

c) Assessed by analysing lung cancer cases and their individually matched controls by conditional logistic regression, conditioning on individual case set, and adjusting for education (in 7 categories) .

d) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 2400 individually matched controls (US cohorts; n=11).

e) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 1189 individually matched controls (European/Australian cohorts; n=5).

f) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 1775 individually matched controls (Asian cohorts; n=4).

g) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 5364 individually matched controls (pooled cohorts; n=20).

h) P for trend assessed by the base 2 logarithm of the circulating concentrations.

Table 4. Odds ratios of lung cancer for circulating concentrations of methionine

			Odds Ratio (95% Confidence Interval)						
LC3 participants	**Quartile (Range)**	**Case/Controls participants**	**Cases compared to matched controls, unadjusted[a] (n=5364/5364)[a]**	**Cases compared to matched controls, adjusted for cotinine and education[b] (n=5364/5364)[b]**	**Never Smokers[c] (n=1327/1327)[c]**	**Former Smokers[c] (n = 1518/1518)[c]**	**Current smokers[b] (n = 2519/2519)[b]**	**Men[b] (n=2908/2908)[b]**	**Women[b] (n=2456/2456)[b]**
	Methionine [µmol/L][d]								
United States (n=11 cohorts)	1 (11.83-22.77)	632/600	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.78-26.16)	617/600	0.97 (0.82 - 1.14)	0.97 (0.82 - 1.14)	1.10 (0.80 - 1.51)	0.82 (0.62 - 1.08)	1.02 (0.78 - 1.35)	0.90 (0.68 - 1.19)	1.00 (0.82 - 1.23)
	3 (26.17-30.47)	591/600	0.92 (0.78 - 1.09)	0.93 (0.78 - 1.10)	1.04 (0.73 - 1.48)	0.78 (0.60 - 1.01)	1.01 (0.75 - 1.36)	0.88 (0.67 - 1.16)	0.95 (0.77 - 1.19)
	4 > 30.47	560/600	0.87 (0.73 - 1.03)	0.87 (0.73 - 1.04)	1.05 (0.68 - 1.62)	0.79 (0.60 - 1.02)	0.84 (0.63 - 1.13)	0.68 (0.52 - 0.89)	1.10 (0.86 - 1.39)
P for trend[h]			0.11	0.12	0.93	0.08	0.36	0.008	0.65
	Methionine [µmol/L][e]								
Europe /Australia (n=5 cohorts)	1 (10.63-22.81)	295/297	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.82-26.29)	316/297	1.08 (0.86 - 1.35)	1.07 (0.85 - 1.35)	1.01 (0.54 - 1.90)	0.88 (0.57 - 1.35)	1.21 (0.88 - 1.65)	1.08 (0.76 - 1.52)	1.07 (0.78 - 1.48)
	3 (26.30-29.91)	282/298	0.94 (0.74 - 1.20)	0.91 (0.71 - 1.16)	0.82 (0.43 - 1.57)	1.06 (0.67 - 1.65)	0.87 (0.62 - 1.20)	0.94 (0.67 - 1.33)	0.87 (0.60 - 1.24)
	4 > 29.91	296/297	0.99 (0.78 - 1.26)	1.00 (0.78 - 1.27)	0.96 (0.49 - 1.89)	1.11 (0.72 - 1.70)	0.95 (0.68 - 1.32)	1.02 (0.74 - 1.42)	0.96 (0.65 - 1.41)
P for trend[h]			0.92	0.92	0.86	0.74	0.83	0.92	0.99
	Methionine [µmol/L][f]								
Asia (n=4 cohorts)	1 (12.32-23.17)	451/443	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.18-27.39)	440/444	0.99 (0.81 - 1.20)	1.02 (0.84 - 1.24)	0.85 (0.63 - 1.16)	0.70 (0.38 - 1.29)	1.32 (0.98 - 1.77)	1.15 (0.89 - 1.49)	0.86 (0.62 - 1.17)
	3 (27.39-32.22)	420/444	0.92 (0.75 - 1.12)	0.93 (0.76 - 1.15)	0.99 (0.71 - 1.38)	0.73 (0.40 - 1.32)	1.00 (0.74 - 1.34)	0.94 (0.73 - 1.22)	1.02 (0.71 - 1.47)
	4 >32.22	464/444	1.02 (0.82 - 1.22)	1.04 (0.85 - 1.26)	1.11 (0.79 - 1.57)	0.59 (0.32 - 1.10)	1.16 (0.88 - 1.53)	1.13 (0.88 - 1.45)	0.86 (0.6 - 1.22)
P for trend[h]			0.82	0.83	0.50	0.17	0.73	0.61	0.70
	Methionine [µmol/L][g]								
All cohorts (n=20 cohorts)	1 (10.63-22.89)	1377/1341	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.90-27.40)	1382/1341	1.00 (0.90 - 1.12)	1.01 (0.91 - 1.13)	0.99 (0.80 - 1.21)	0.88 (0.72 - 1.09)	1.12 (0.95 - 1.32)	1.02 (0.87 - 1.21)	1.00 (0.86 - 1.16)
	3 (27.40-31.00)	1303/1341	0.93 (0.83 - 1.05)	0.94 (0.84 - 1.05)	0.98 (0.78 - 1.22)	0.80 (0.65 - 0.99)	1.01 (0.84 - 1.20)	0.89 (0.76 - 1.05)	1.00 (0.85 - 1.17)
	4 > 31.01	1302/1341	0.93 (0.83 - 1.04)	0.95 (0.85 - 1.07)	1.04 (0.81 - 1.33)	0.83 (0.67 - 1.03)	0.99 (0.83 - 1.17)	0.92 (0.79 - 1.08)	0.99 (0.83 - 1.18)
P for trend[h]			0.22	0.42	0.52	0.07	0.69	0.22	0.88

a) Assessed by analysing lung cancer cases and their individually matched controls by conditional logistic regression, conditioning on individual case set.

b) Assessed by analysing lung cancer cases and their individually matched controls by conditional logistic regression, conditioning on individual case set, and adjusting for circulating cotinine (in quartiles) and education (in 7 categories) .

c) Assessed by analysing lung cancer cases and their individually matched controls by conditional logistic regression, conditioning on individual case set, and adjusting for education (in 7 categories) .

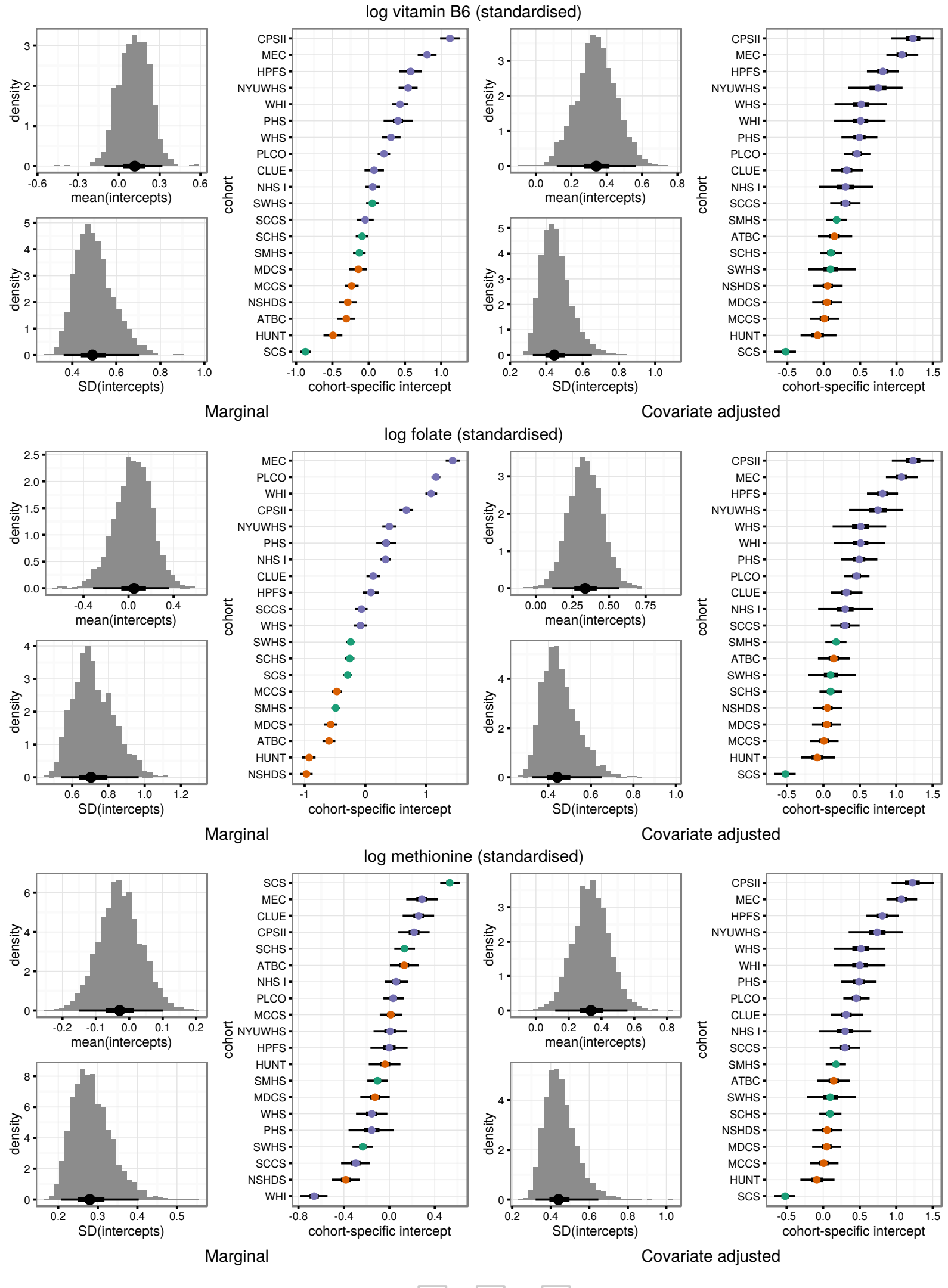
d) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 2400 individually matched controls (US cohorts; n=11).

e) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 1189 individually matched controls (European/Australian cohorts; n=5).

f) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 1775 individually matched controls (Asian cohorts; n=4).

g) Quartile cut-off points were determined based on the blood concentration distribution of each biomarker for 5364 individually matched controls (pooled cohorts; n=20).

h) P for trend assessed by the base 2 logarithm of the circulating concentrations.

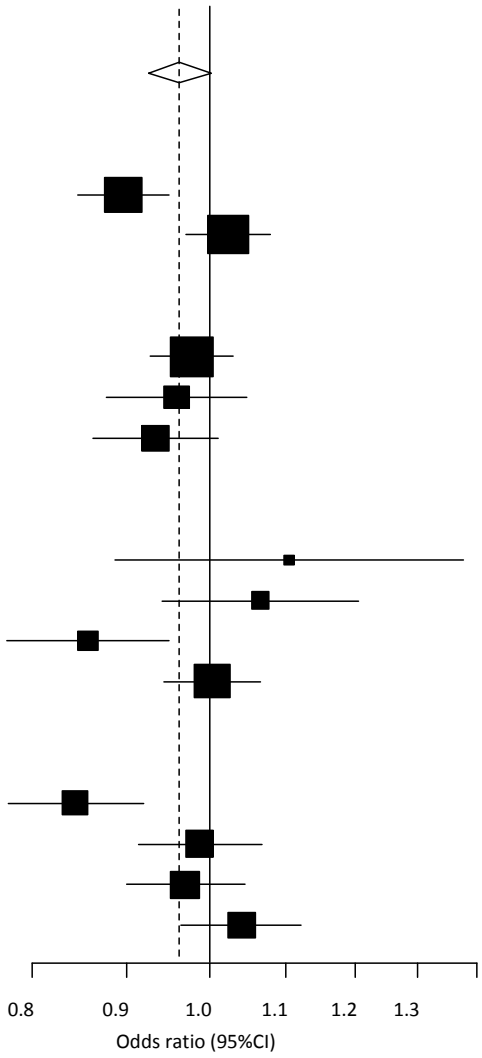

log vitamin B6 (standardised)
density
mean(intercepts)
SD(intercepts)
cohort
cohort-specific intercept
CPSII
MEC
HPFS
NYUWHS
WHI
PHS
WHS
PLCO
CLUE
NHS I
SWHS
SCCS
SCHS
SMHS
MDCS
MCCS
NSHDS
ATBC
HUNT
SCS
Marginal
Covariate adjusted
log folate (standardised)
Marginal
Covariate adjusted
log methionine (standardised)
Marginal
Covariate adjusted
continent
ASIA
EU/AU
US

Supplementary Files:

Supplementary Figures

Supplementary Figure 1. Forest plot showing stratified ORs of lung cancer for log2- of circulating vitamin B6 overall.

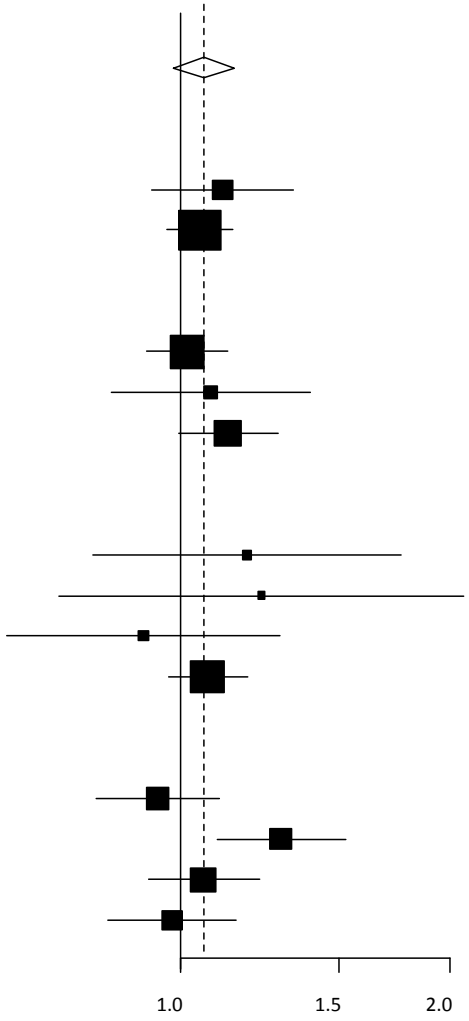
study group	Cases	Control	OR[a]	95%CI
Overall log2-Vitamin B6	5364	5364	0.96	0.93-1.00
By gender (p-heterogeneity= 0.001)[b]				
Men	2908	2908	0.90	0.85-0.95
Women	2456	2456	1.02	0.97-1.08
By Continent (p-heterogeneity= 0.83)[b]				
United States	2400	2400	0.98	0.93-1.03
Europe/Australia	1189	1189	0.96	0.88-1.05
Asia	1775	1775	0.93	0.86-1.01
By histology (p-heterogeneity= 0.03)[b]				
Large cell carcinoma	174	5364	1.11	0.89-1.38
Small cell carcinoma	492	5364	1.07	0.94-1.21
Squamous cell carcinoma	836	5364	0.86	0.77-0.95
Adenocarcinoma	2056	5364	1.00	0.94-1.07
By time from blood draw to diagnosis (p-heterogeneity=0.001)[b]				
1-35 months	1086	5364	0.85	0.78-0.92
36-72 months	1256	5364	0.99	0.91-1.07
73-120 months	1445	5364	0.97	0.90-1.05
>120 months	1577	5364	1.04	0.97-1.12



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates.
[b] $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 2. Forest plot showing stratified ORs of lung cancer for log2- of circulating vitamin B6 among never smokers.

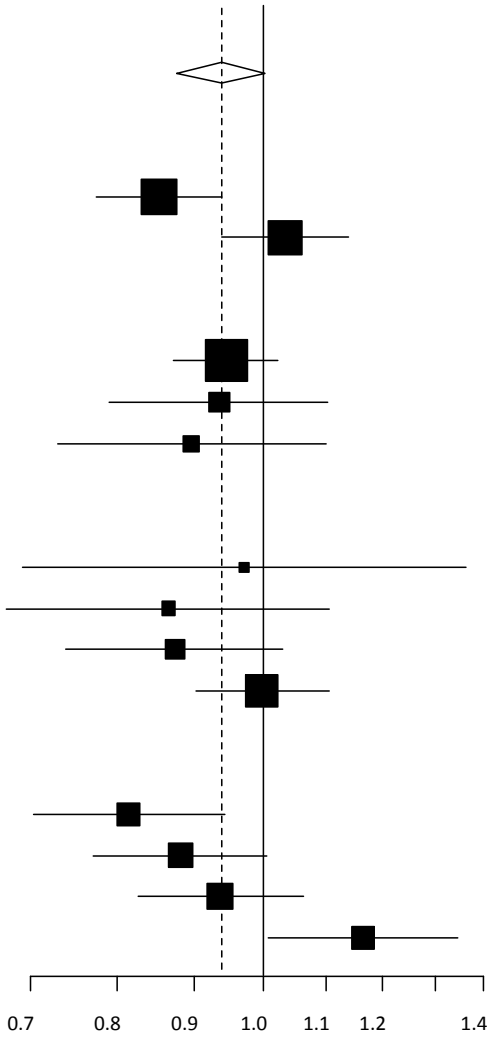
study group	Cases	Controls	OR[a]	95%CI
Never smokers log2-Vitamin B6	1327	1327	1.06	0.98-1.15
By gender (p-heterogeneity= 0.72)[b]				
Men	292	292	1.11	0.93-1.34
Women	1035	1035	1.05	0.96-1.14
By Continent (p-heterogeneity= 0.49)[b]				
United States	569	569	1.02	0.92-1.13
Europe/Australia	156	156	1.08	0.83-1.40
Asia	602	602	1.13	0.99-1.29
By histology (p-heterogeneity= 0.71)[b]				
Large cell carcinoma	29	1327	1.19	0.80-1.76
Small cell carcinoma	27	1327	1.23	0.73-2.08
Squamous cell carcinoma	59	1327	0.91	0.64-1.29
Adenocarcinoma	731	1327	1.07	0.97-1.19
By time from blood draw to diagnosis (p-heterogeneity=0.02)[b]				
1-35 months	263	1327	0.94	0.81-1.10
36-72 months	321	1327	1.30	1.10-1.53
73-120 months	393	1327	1.06	0.92-1.22
>120 months	350	1327	0.98	0.83-1.15



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates.
[b] $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 3. Forest plot showing stratified ORs of lung cancer for log2- of circulating vitamin B6 among former smokers.

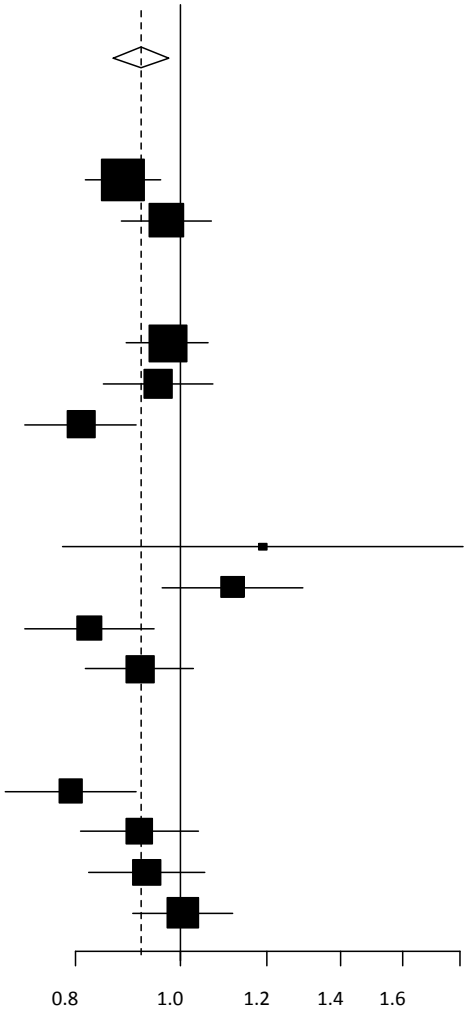
study group	Cases	Controls	OR[a]	95%CI
Former smokers log2-Vitamin B6	1518	1518	0.94	0.88-1.00
By gender (p-heterogeneity= 0.005)[b]				
Men	898	898	0.85	0.77-0.94
Women	620	620	1.03	0.94-1.14
By Continent (p-heterogeneity= 0.86)[b]				
United States	1007	1007	0.94	0.87-1.02
Europe/Australia	335	335	0.93	0.79-1.10
Asia	176	176	0.90	0.73-1.10
By histology (p-heterogeneity= 0.72)[b]				
Large cell carcinoma	70	1518	0.97	0.69-1.36
Small cell carcinoma	135	1518	0.86	0.67-1.11
Squamous cell carcinoma	246	1518	0.87	0.74-1.03
Adenocarcinoma	642	1518	1.00	0.90-1.11
By time from blood draw to diagnosis (p-heterogeneity=0.003)[b]				
1-35 months	344	1518	0.81	0.70-0.94
36-72 months	368	1518	0.88	0.77-1.01
73-120 months	419	1518	0.94	0.82-1.06
>120 months	387	1518	1.16	1.01-1.35



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates.
[b] $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 4. Forest plot showing stratified ORs of lung cancer for log2- of circulating vitamin B6 among current smokers.

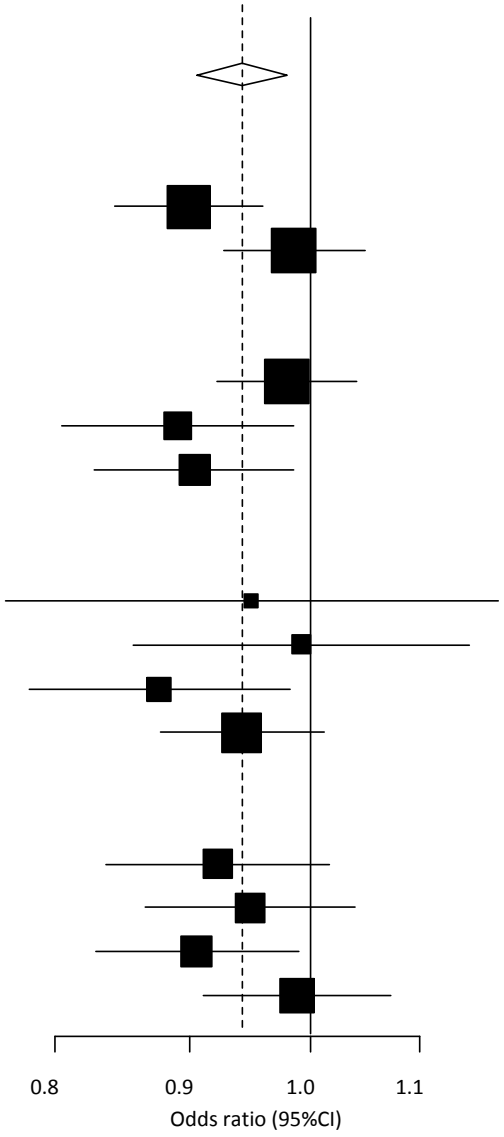
study group	Cases	Controls	OR[a]	95%CI
Current smokers log2-Vitamin B6	2519	2519	0.92	0.87-0.98
By gender (p-heterogeneity= 0.11)[b]				
Men	1718	1718	0.88	0.82-0.96
Women	801	801	0.97	0.88-1.07
By Continent (p-heterogeneity= 0.04)[b]				
United States	824	824	0.97	0.89-1.06
Europe/Australia	698	698	0.95	0.85-1.07
Asia	997	997	0.81	0.72-0.91
By histology (p-heterogeneity= 0.01)[b]				
Large cell carcinoma	75	2519	1.19	0.78-1.82
Small cell carcinoma	330	2519	1.12	0.96-1.30
Squamous cell carcinoma	531	2519	0.82	0.72-0.95
Adenocarcinoma	683	2519	0.92	0.82-1.03
By time from blood draw to diagnosis (p-heterogeneity=0.07)[b]				
1-35 months	479	2519	0.79	0.69-0.91
36-72 months	567	2519	0.92	0.81-1.04
73-120 months	633	2519	0.93	0.82-1.05
>120 months	840	2519	1.00	0.90-1.12



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates.
[b] $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 5. Forest plot showing stratified ORs of lung cancer for log2- of circulating folate overall.

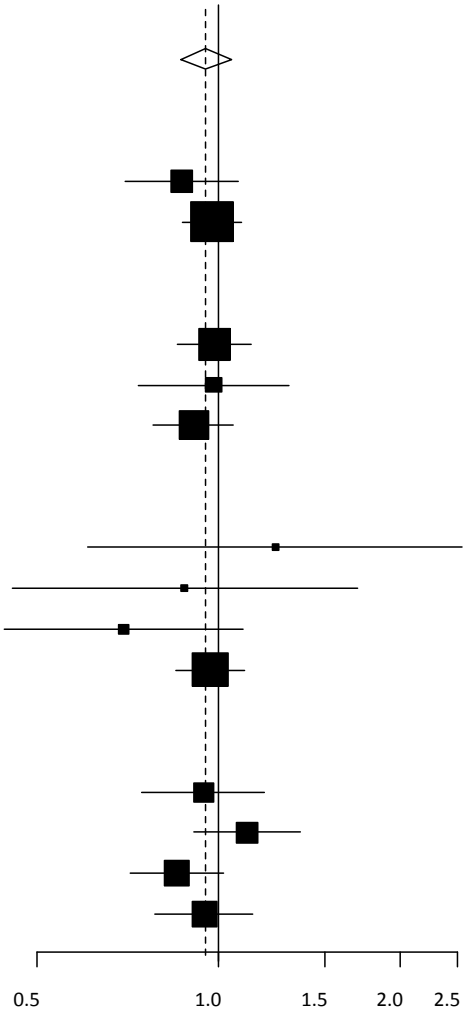
study group	Cases	Controls	OR[a]	95%CI
Overall log2-Folate	5364	5364	0.94	0.90-0.97
By gender (p-heterogeneity= 0.05)[b]				
Men	2908	2908	0.90	0.84-0.96
Women	2456	2456	0.99	0.93-1.05
By Continent (p-heterogeneity= 0.11)[b]				
United States	2400	2400	0.98	0.92-1.04
Europe/Australia	1189	1189	0.89	0.80-0.99
Asia	1775	1775	0.90	0.83-0.99
By histology (p-heterogeneity= 0.65)[b]				
Large cell carcinoma	174	5364	0.95	0.77-1.18
Small cell carcinoma	492	5364	0.99	0.86-1.15
Squamous cell carcinoma	836	5364	0.88	0.78-0.98
Adenocarcinoma	2056	5364	0.94	0.88-1.01
By time from blood draw to diagnosis (p-heterogeneity=0.53)[b]				
1-35 months	1086	5364	0.92	0.84-1.02
36-72 months	1256	5364	0.95	0.87-1.04
73-120 months	1445	5364	0.91	0.83-0.99
>120 months	1577	5364	0.99	0.91-1.07



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates.
[b] $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 6. Forest plot showing stratified ORs of lung cancer for log2- of circulating folate among never smokers.

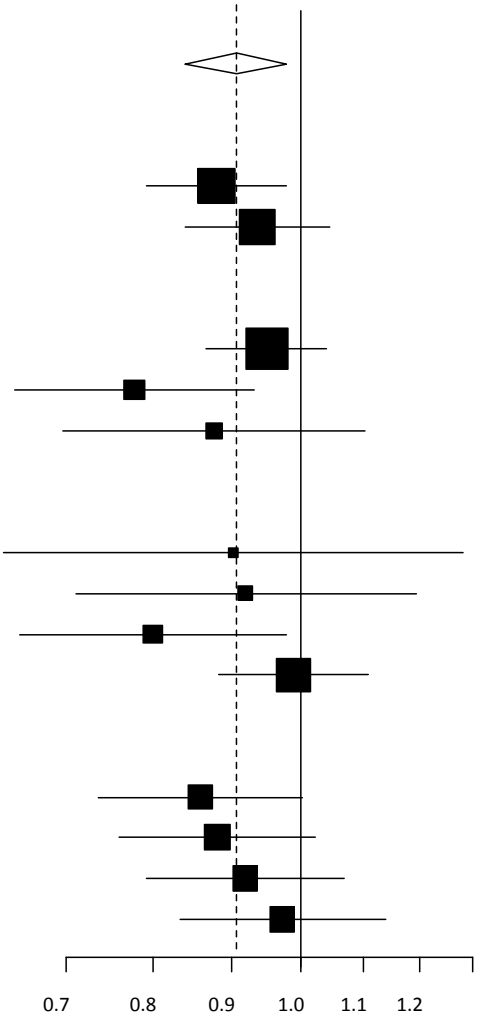
study group	Cases	Controls	OR[a]	95%CI
Never smokers log2-Folate	1327	1327	0.95	0.86-1.05
By gender (p-heterogeneity= 0.33)[b]				
Men	292	292	0.87	0.70-1.08
Women	1035	1035	0.98	0.88-1.09
By Continent (p-heterogeneity= 0.72)[b]				
United States	569	569	0.99	0.86-1.14
Europe/Australia	156	156	0.98	0.74-1.31
Asia	602	602	0.91	0.78-1.06
By histology (p-heterogeneity= 0.50)[b]				
Large cell carcinoma	29	1327	1.24	0.61-2.55
Small cell carcinoma	27	1327	0.88	0.45-1.70
Squamous cell carcinoma	59	1327	0.70	0.44-1.10
Adenocarcinoma	731	1327	0.97	0.85-1.11
By time from blood draw to diagnosis (p-heterogeneity=0.27)[b]				
1-35 months	263	1327	0.94	0.75-1.19
36-72 months	321	1327	1.12	0.91-1.37
73-120 months	393	1327	0.85	0.71-1.02
>120 months	350	1327	0.95	0.79-1.14



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates.
[b] $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 7. Forest plot showing stratified ORs of lung cancer for log2- of circulating folate among former smokers.

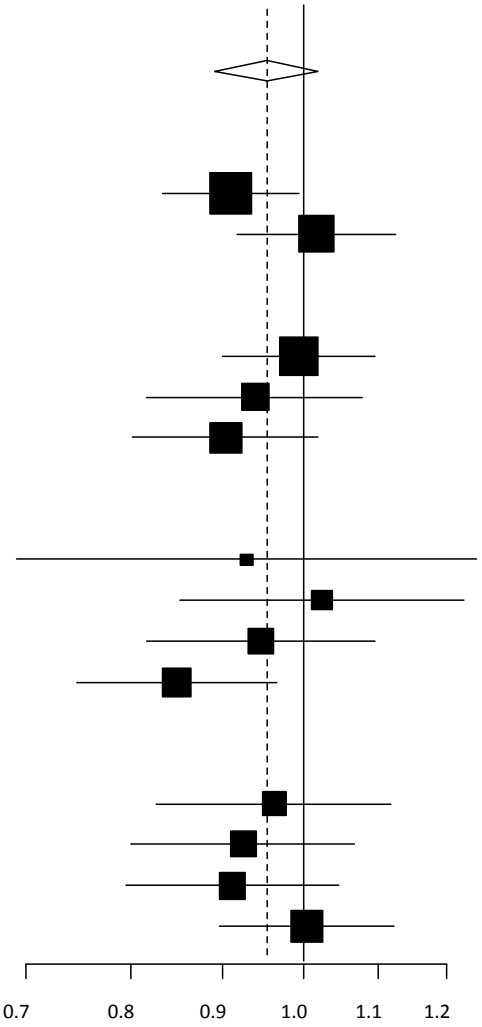
study group	Cases	Controls	OR[a]	95%CI
Former smokers log2-Folate	1518	1518	0.91	0.84-0.98
By gender (p-heterogeneity= 0.41)[b]				
Men	898	898	0.88	0.79-0.98
Women	620	620	0.94	0.84-1.05
By Continent (p-heterogeneity= 0.14)[b]				
United States	1007	1007	0.95	0.87-1.04
Europe/Australia	335	335	0.78	0.65-0.93
Asia	176	176	0.88	0.70-1.10
By histology (p-heterogeneity= 0.37)[b]				
Large cell carcinoma	70	1518	0.90	0.64-1.28
Small cell carcinoma	135	1518	0.92	0.71-1.19
Squamous cell carcinoma	246	1518	0.80	0.65-0.98
Adenocarcinoma	642	1518	0.99	0.88-1.11
By time from blood draw to diagnosis (p-heterogeneity=0.71)[b]				
1-35 months	344	1518	0.86	0.74-1.00
36-72 months	368	1518	0.88	0.76-1.02
73-120 months	419	1518	0.92	0.79-1.07
>120 months	387	1518	0.97	0.83-1.14



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates.
[b] $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 8. Forest plot showing stratified ORs of lung cancer for log2- of circulating folate among current smokers.

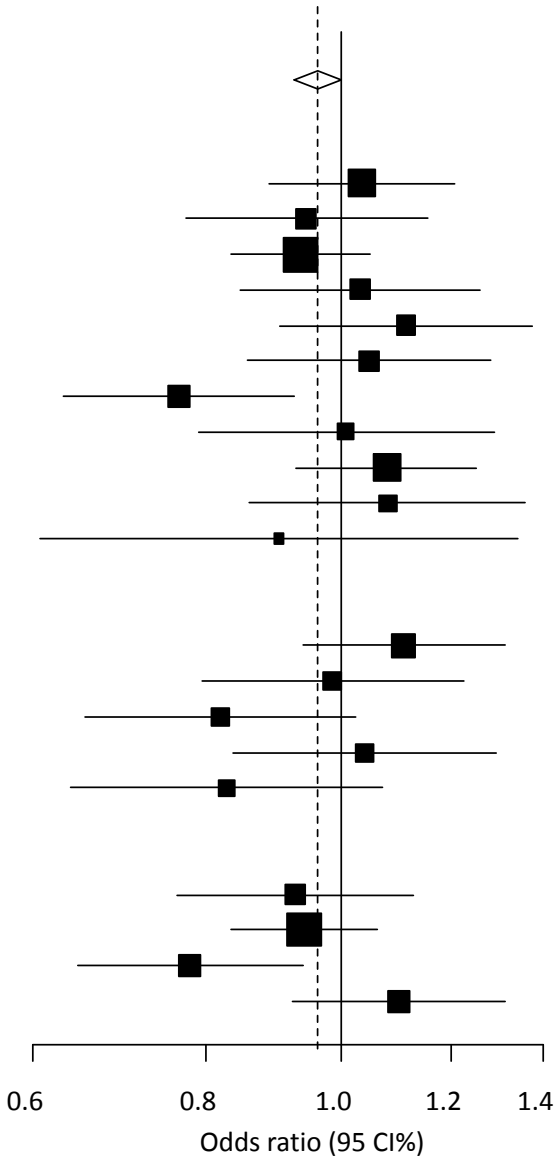
study group	Cases	Controls	OR[a]	95%CI
Current smokers log2-Folate	2519	2519	0.95	0.89-1.02
By gender (p-heterogeneity= 0.11)[b]				
Men	1718	1718	0.91	0.83-0.99
Women	801	801	1.02	0.92-1.12
By Continent (p-heterogeneity= 0.46)[b]				
United States	824	824	0.99	0.90-1.09
Europe/Australia	698	698	0.94	0.82-1.08
Asia	997	997	0.90	0.80-1.02
By histology (p-heterogeneity= 0.40)[b]				
Large cell carcinoma	75	2519	0.93	0.69-1.25
Small cell carcinoma	330	2519	1.02	0.85-1.23
Squamous cell carcinoma	531	2519	0.95	0.82-1.09
Adenocarcinoma	683	2519	0.85	0.75-0.97
By time from blood draw to diagnosis (p-heterogeneity=0.67)[b]				
1-35 months	479	2519	0.96	0.83-1.12
36-72 months	567	2519	0.92	0.80-1.07
73-120 months	633	2519	0.91	0.80-1.04
>120 months	840	2519	1.00	0.90-1.12



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates.
[b] $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 9. Forest plot showing stratified ORs of lung cancer for log2 of circulating vitamin B6 by LC3 cohorts.

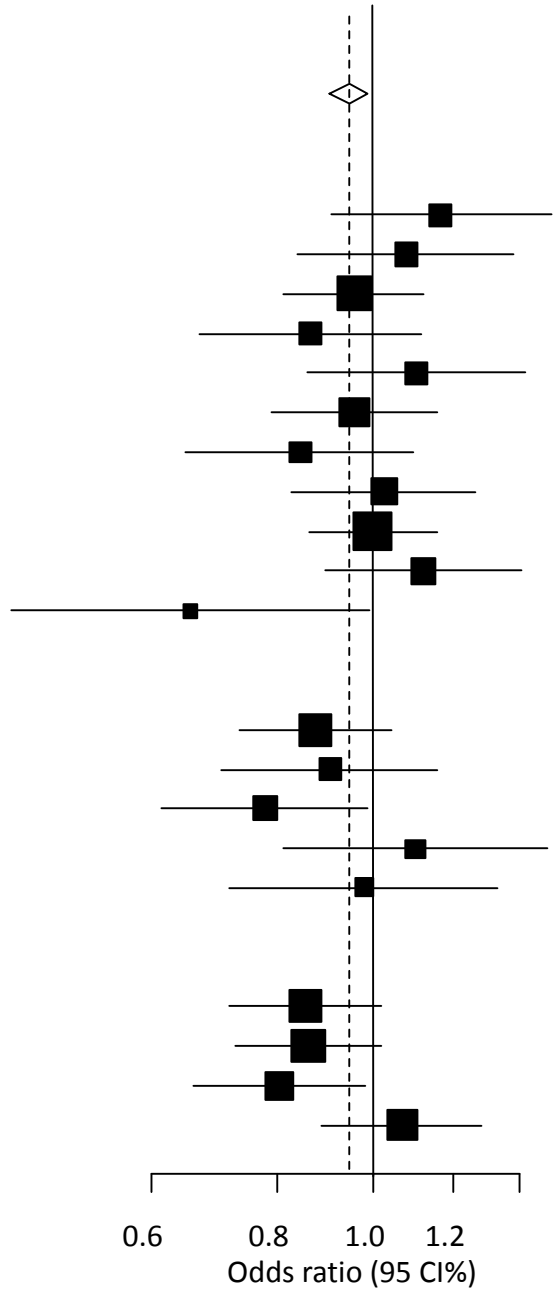
study group	Cases	Controls	OR[a]	95%CI
Overall log2-Vitamin B6	5364	5364	0.96	0.93-1.00
US cohorts				
WHI	241	241	1.04	0.89-1.21
SCCS	226	226	0.95	0.77-1.16
PLCO	450	450	0.94	0.83-1.05
NYUWHS	171	171	1.03	0.85-1.26
CPS	182	182	1.12	0.90-1.38
CLUE	191	191	1.05	0.86-1.28
MEC	174	174	0.77	0.63-0.93
WHS	184	184	1.01	0.79-1.29
NHS	345	345	1.08	0.93-1.25
HPFS	155	155	1.08	0.86-1.36
PHS	81	81	0.90	0.61-1.34
EU/AU cohorts				
MCCS	354	354	1.11	0.94-1.32
MDCS	198	198	0.99	0.80-1.23
NSHDS	244	244	0.82	0.66-1.02
ATBC	200	200	1.04	0.84-1.30
HUNT	193	193	0.83	0.64-1.07
Asian cohorts				
SCS	513	513	0.93	0.76-1.13
SCHS	422	422	0.94	0.83-1.06
SMHS	421	421	0.78	0.65-0.94
SWHS	419	419	1.10	0.92-1.32



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates. $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 10. Forest plot showing stratified ORs of lung cancer for log2 of circulating folate by LC3 cohorts.

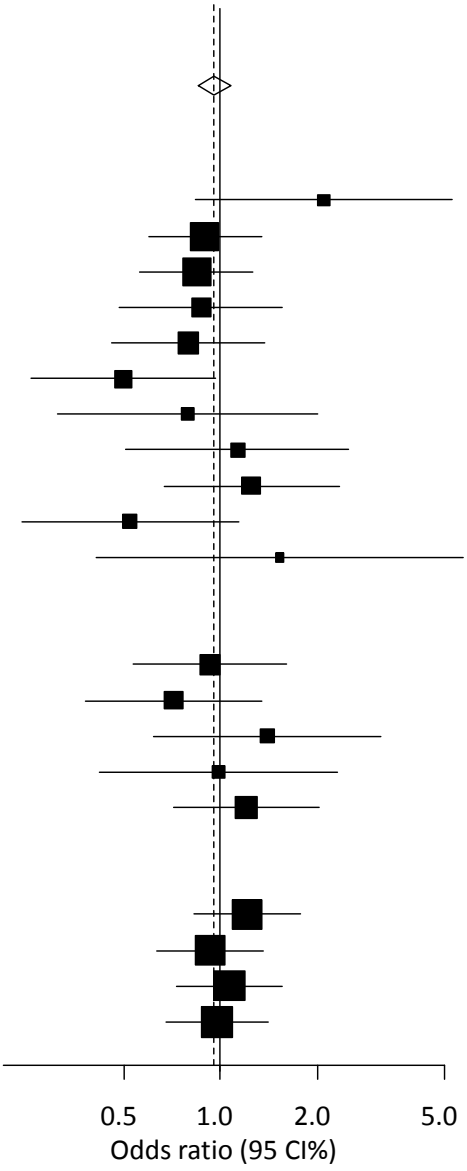
study group	Cases	Controls	OR[a]	95%CI
Overall log2-Folate	5364	5364	0.94	0.90-0.97
US cohorts				
WHI	241	241	1.17	0.91-1.50
SCCS	226	226	1.08	0.84-1.38
PLCO	450	450	0.95	0.81-1.12
NYUWHS	171	171	0.86	0.67-1.12
CPS	182	182	1.10	0.86-1.41
CLUE	191	191	0.96	0.79-1.16
MEC	174	174	0.84	0.65-1.10
WHS	184	184	1.02	0.83-1.26
NHS	345	345	1.00	0.86-1.16
HPFS	155	155	1.12	0.90-1.40
PHS	81	81	0.66	0.43-0.99
EU/AU cohorts				
MCCS	354	354	0.87	0.73-1.04
MDCS	198	198	0.90	0.70-1.16
NSHDS	244	244	0.78	0.61-0.99
ATBC	200	200	1.10	0.81-1.49
HUNT	193	193	0.98	0.72-1.33
Asian cohorts				
SCS	513	513	0.85	0.72-1.02
SCHS	422	422	0.86	0.73-1.02
SMHS	421	421	0.81	0.66-0.98
SWHS	419	419	1.07	0.89-1.28



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates. $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

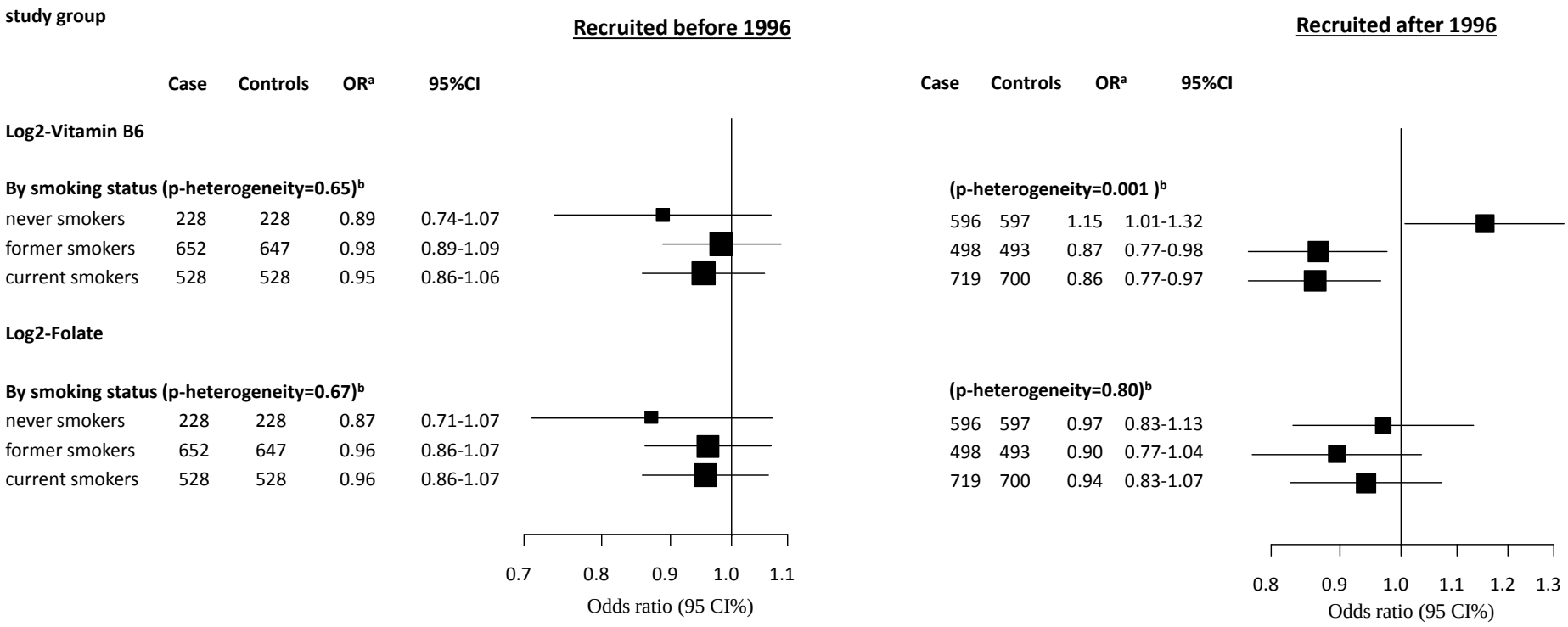
Supplementary Figure 11. Forest plot showing stratified ORs of lung cancer for log2 of circulating methionine by LC3 cohorts.

study group	Cases	Controls	OR[a]	95%CI
Overall log2-Methionine	5364	5364	0.96	0.85-1.08
US cohorts				
WHI	241	241	2.10	0.83-5.30
SCCS	226	226	0.90	0.60-1.34
PLCO	450	450	0.84	0.56-1.26
NYUWHS	171	171	0.87	0.49-1.56
CPS	182	182	0.79	0.46-1.37
CLUE	191	191	0.50	0.26-0.97
MEC	174	174	0.79	0.31-2.01
WHS	184	184	1.13	0.51-2.51
NHS	345	345	1.25	0.67-2.34
HPFS	155	155	0.52	0.24-1.14
PHS	81	81	1.53	0.41-5.71
EU/AU cohorts				
MCCS	354	354	0.92	0.53-1.60
MDCS	198	198	0.71	0.38-1.34
NSHDS	244	244	1.40	0.62-3.18
ATBC	200	200	0.99	0.42-2.33
HUNT	193	193	1.20	0.71-2.03
Asian cohorts				
SCS	513	513	1.21	0.82-1.77
SCHS	422	422	0.93	0.63-1.37
SMHS	421	421	1.07	0.73-1.55
SWHS	419	419	0.97	0.68-1.41



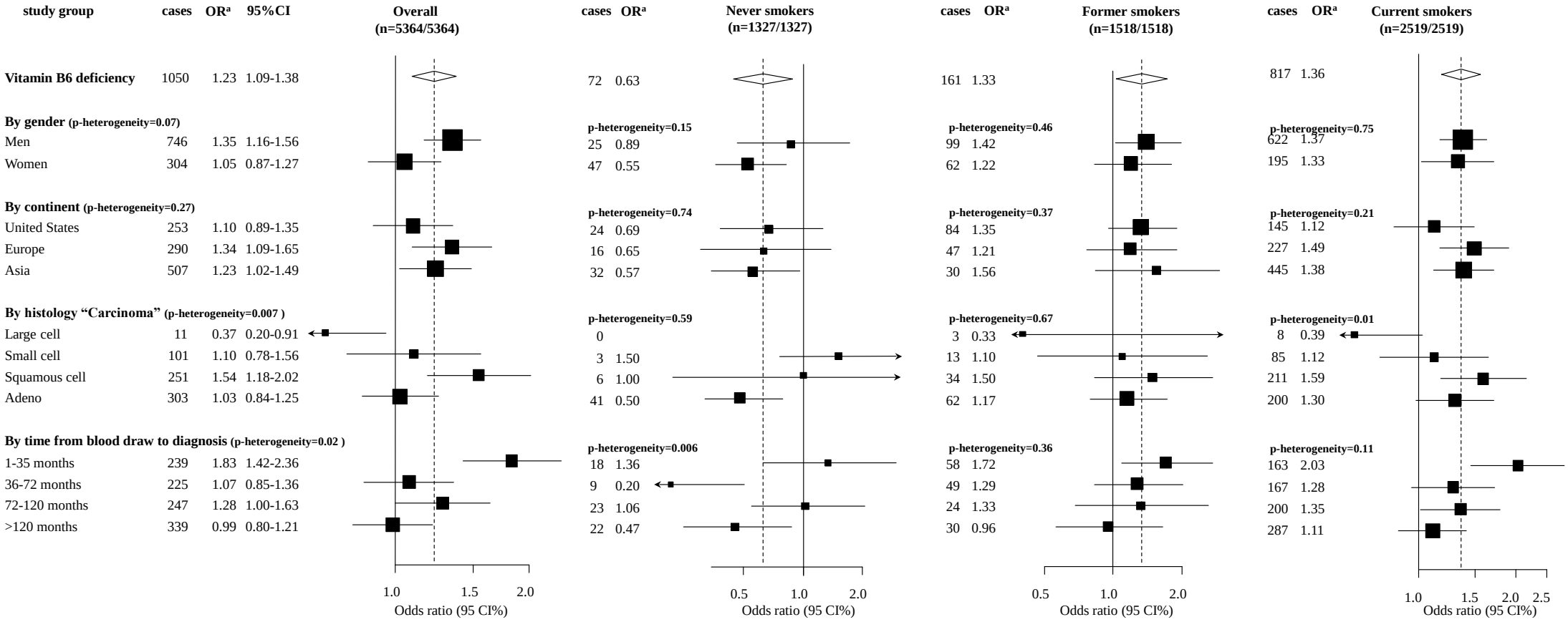
[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates. $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 12. Forest plot showing stratified ORs of lung cancer for log2 of circulating vitamin B6 and folate by period of recruitment and by smoking status in US cohorts (n=2400 cases and matched controls).



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% Cis; the size of the black squares are proportional to the inverse variance of logistic regression estimates. [b] $P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

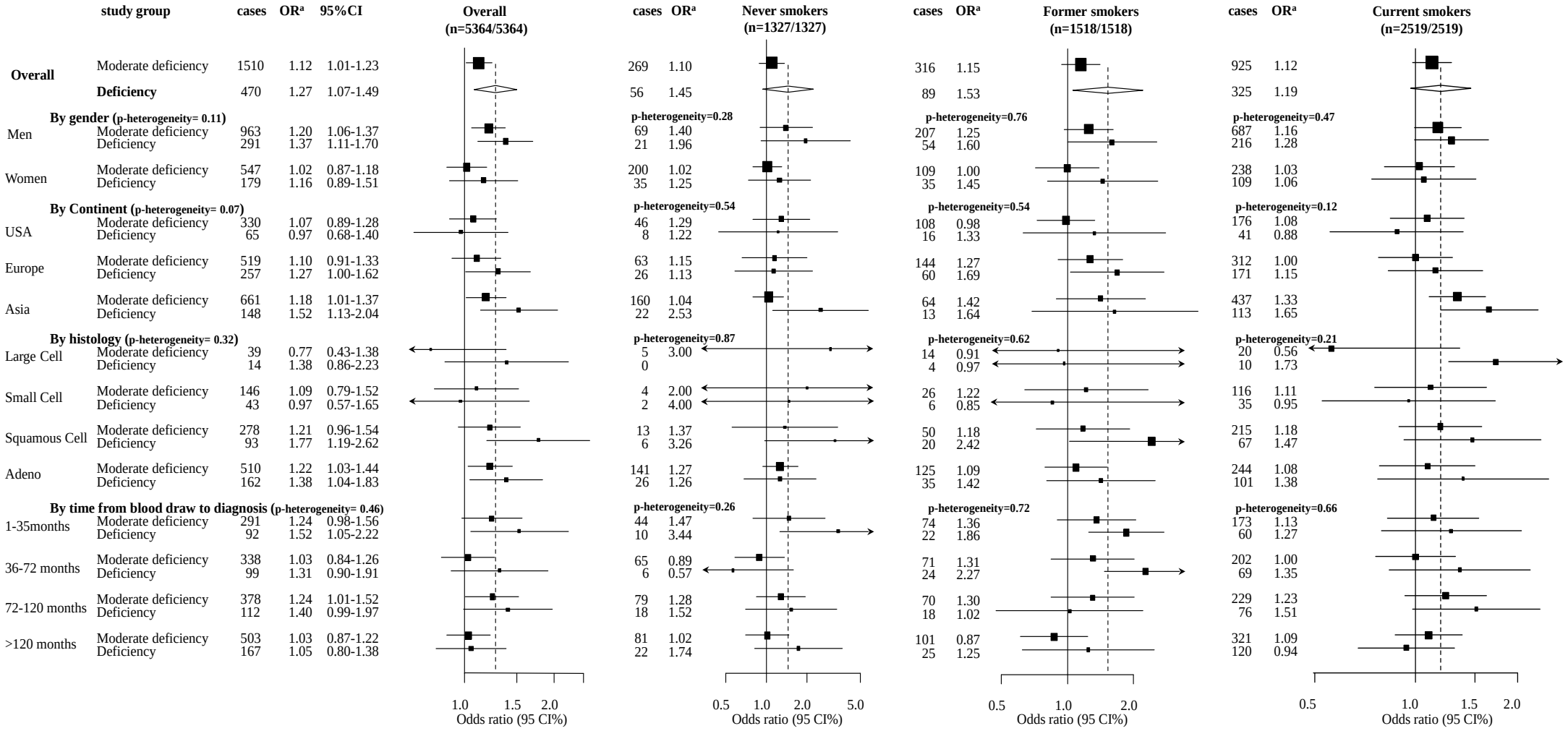
Supplementary Figure 13. Forest plot showing stratified ORs of lung cancer for B6 deficiency overall and by smoking status.



[a]ORs were assessed by conditional logistic regression, and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% Cis; the size of the black squares are proportional to the inverse variance of logistic regression estimates.

Cases listed are those from the group presenting a B6 deficiency. Number of cases within the reference group is not listed. Subjects were considered vitamin B6 deficient if their circulating concentrations were less than 20nmol/L and normal if above or equal to 20nmol/L.

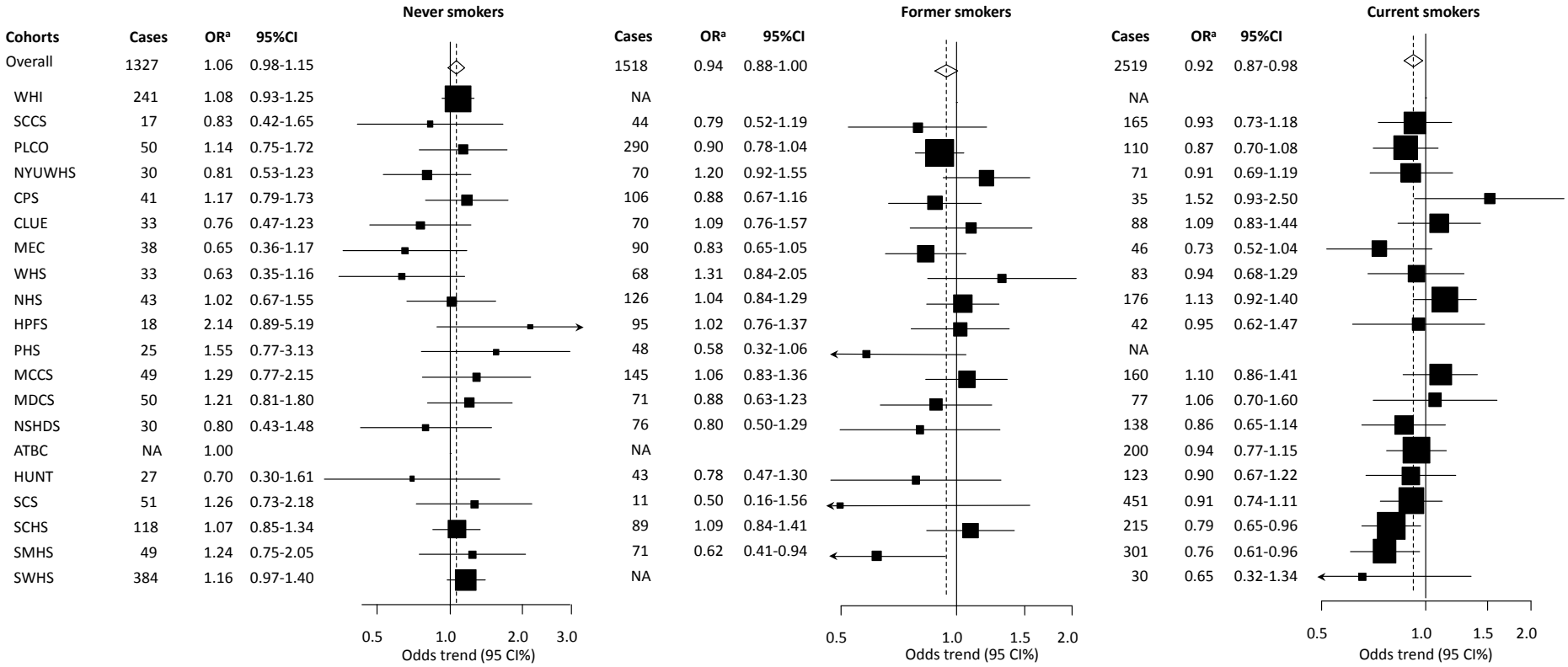
Supplementary Figure 14. Forest plot showing stratified ORs of lung cancer for Folate deficiency overall and by smoking status.



[a]ORs were assessed by conditional logistic regression, and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% Cis; the size of the black squares are proportional to the inverse variance of logistic regression estimates.

Cases listed are those from the group presenting a folate deficiency or a moderate deficiency. Number of cases within the reference group is not listed. For folate deficiency, subjects were classified in three categories; deficient (folate <7 nmol/L), moderately deficient (folate 7-13 nmol/L), or normal (folate ≥ 13 nmol/L).

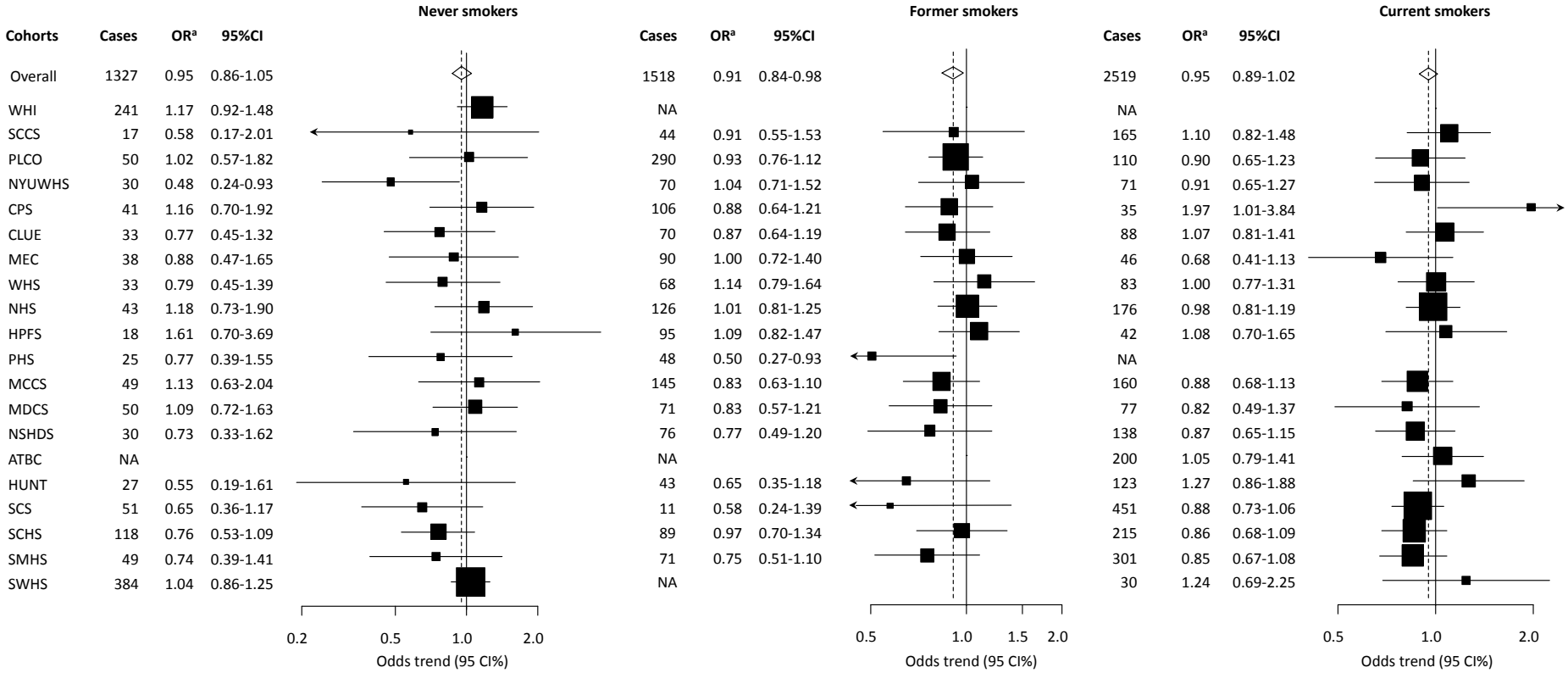
Supplementary Figure 15. Forest plot showing stratified ORs of lung cancer for log2- of circulating vitamin B6 for each cohort participants by smoking status.



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates. Sample sizes lower than 10 were classified as not available (NA).

$P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Figure 16. Forest plot showing stratified ORs of lung cancer for log2- of circulating folate for each cohort participants by smoking status.



[a]ORs were assessed by conditional logistic regression by including the base 2 logarithm of circulating concentrations (ORs indicate relative risks of a doubling in circulating concentrations), and where relevant adjusted for cotinine and education; the black squares and horizontal lines indicate the ORs and 95% CIs; the size of the black squares is proportional to the inverse variance of the logistic regression estimates. Sample sizes lower than 10 were classified as not available (NA).

$P_{heterogeneity}$ indicates results of chi-square test assessing the null hypothesis of ORs being identical.

Supplementary Tables

Supplementary Table 1. Within-batch % coefficients of variation for methionine, vitamin B6 and folate, respectively.

Batch	Folate	PLP	Methionine	Batch	Folate	PLP	Methionine	Batch	Folate	PLP	Methionine	Batch	Folate	PLP	Methionine
1	13.7	3.8	3.3	41	4.9	4.6	1.5	81	3.7	3.5	3.0	121	4.8	3.2	1.5
2	7.4	4.3	2.6	42	8.5	4.9	3.2	82	8.8	3.3	2.5	122	5.9	5.8	3.0
3	1.2	2.2	3.8	43	5.3	5.0	2.3	83	3.6	4.0	2.7	123	0.7	4.0	2.5
4	8.5	2.5	4.0	44	1.9	3.1	1.2	84	7.6	3.3	2.3	124	12.6	4.2	2.4
5	7.8	3.4	3.5	45	2.6	2.7	2.3	85	3.5	5.7	4.9	125	11.8	2.4	2.0
6	1.3	3.9	5.8	46	0.1	3.5	3.8	86	12.9	4.6	6.4	126	2.1	2.5	2.2
7	2.3	3.2	2.1	47	13.2	6.4	1.9	87	8.7	6.9	1.7	127	2.7	3.7	2.8
8	2.8	4.6	1.6	48	12.6	2.2	2.0	88	0.1	2.6	1.5	128	3.0	6.0	6.0
9	3.3	5.0	3.7	49	2.0	2.2	1.6	89	16.3	4.5	2.7	129	0.5	3.7	3.8
10	5.3	4.6	4.4	50	13.7	2.6	1.5	90	11.1	2.5	2.4	130	4.1	2.5	1.7
11	0.2	3.7	2.7	51	10.7	3.2	3.4	91	5.8	4.0	3.4	131	11.9	6.1	2.7
12	4.4	3.8	3.9	52	7.7	2.9	1.7	92	7.9	5.9	2.5	132	0.5	6.6	3.4
13	0.5	10.3	1.6	53	2.3	5.1	1.5	93	7.8	5.5	3.5	133	6.7	6.8	2.2
14	0.7	4.8	4.1	54	5.9	2.2	1.5	94	3.5	4.9	3.3	134	5.7	4.3	4.8
15	3.3	6.6	2.1	55	9.5	2.9	2.9	95	14.5	4.2	5.9	135	15.1	7.3	2.7
16	5.3	10.7	4.9	56	18.0	3.6	2.5	96	8.8	2.9	2.4	136	14.2	5.3	2.9
17	9.6	12.0	3.4	57	5.2	4.9	2.3	97	2.2	6.9	2.5	137	0.6	4.6	1.6
18	0.3	3.2	1.4	58	5.3	4.8	3.2	98	0.9	1.5	1.8	138	10.5	6.2	2.5
19	1.7	2.9	3.5	59	1.9	6.1	2.2	99	4.2	5.6	2.3	139	5.8	4.5	2.4
20	0.6	10.2	4.8	60	3.7	4.7	2.4	100	11.4	5.7	2.7	140	0.2	5.1	1.6
21	4.6	9.9	2.7	61	12.0	9.5	2.8	101	10.1	5.6	2.2	141	16.7	5.5	0.7
22	2.9	6.4	3.9	62	8.4	7.1	2.8	102	3.8	8.0	2.3	142	8.0	6.8	2.7
23	1.1	5.9	1.7	63	5.5	4.0	1.1	103	2.5	3.9	2.9	143	3.9	7.6	2.2
24	1.8	3.6	4.3	64	2.6	5.8	2.0	104	1.2	6.1	3.5	144	4.3	5.4	2.7
25	2.3	13.6	4.9	65	8.4	5.9	3.9	105	3.7	7.5	2.6	145	5.0	8.7	2.5
26	4.3	3.9	2.7	66	4.1	1.9	3.9	106	3.7	6.1	3.9	146	8.2	5.8	2.3
27	4.1	11.3	2.9	67	0.2	2.2	4.2	107	3.0	6.4	1.2	147	2.0	15.2	1.6
28	4.4	8.5	3.6	68	14.4	3.0	3.9	108	0.1	6.4	2.0	148	5.4	5.4	3.2
29	4.3	7.6	4.1	69	13.7	2.2	1.9	109	4.5	3.5	3.9	149	2.8	7.4	1.9
30	6.6	14.6	8.2	70	13.0	4.1	2.6	110	3.1	3.1	4.3	150	4.4	11.7	4.2
31	2.9	8.4	2.9	71	9.4	7.5	1.5	111	1.7	3.0	1.7	151	4.5	7.6	4.1
32	13.1	5.1	1.6	72	5.9	4.3	2.1	112	10.0	7.5	3.2	152	2.5	17.6	2.6
33	2.0	8.9	1.2	73	9.8	2.1	2.3	113	3.7	5.3	1.5				
34	1.9	5.7	3.5	74	13.8	2.4	2.6	114	10.6	2.0	2.9				
35	1.7	4.6	2.1	75	10.1	5.8	2.0	115	8.9	3.1	4.3				
36	11.2	2.5	2.3	76	0.8	5.5	3.8	116	7.1	7.2	3.4				
37	0.6	3.6	2.7	77	4.8	2.8	1.9	117	3.7	5.1	1.3				
38	9.5	1.0	2.4	78	8.8	3.1	1.6	118	11.1	4.9	2.6				
39	1.4	6.3	1.8	79	10.6	4.5	2.6	119	12.3	2.6	3.2				
40	9.7	6.8	2.1	80	5.6	4.2	4.1	120	5.7	7.3	2.3				

Supplementary Table 1. Odds ratios of lung cancer for circulating levels of vitamin B6, folate and methionine among ever smoking participants in US

		US cohorts				
				Model 1[a]	Model 2[b]	Model 3[c]
		Cases	Controls	OR (95%CI)	OR (95%CI)	OR (95%CI)
All ever smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (5.08-29.89)	487	455	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (29.90-49.90)	386	387	0.96 (0.79 - 1.17)	0.98 (0.8 - 1.2)	0.96 (0.78 - 1.18)
	3 (49.91-90.05)	328	343	0.93 (0.76 - 1.14)	0.95 (0.76 - 1.17)	0.91 (0.74 - 1.13)
	4 > 90.05	322	338	0.92 (0.74 - 1.14)	0.97 (0.77 - 1.21)	0.97 (0.78 - 1.22)
	P for trend[e]			0.32	0.61	0.57
	Folate (Vitamin B9) [nmol/L][d]					
	1 (1.70-17.45)	450	429	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (17.46-33.35)	369	388	0.93 (0.76 - 1.14)	0.96 (0.78 - 1.19)	0.95 (0.77 - 1.17)
	3 (33.36-59.10)	353	350	0.98 (0.78 - 1.22)	1.01 (0.8 - 1.27)	1 (0.8 - 1.26)
	4 > 59.10	351	356	0.98 (0.77 - 1.25)	1.03 (0.81 - 1.32)	1.03 (0.81 - 1.32)
	P for trend[e]	.	.	0.75	0.96	0.97
	Methionine [µmol/L][d]					
	1 (11.83-22.77)	385	363	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.78-26.16)	375	367	0.94 (0.76 - 1.16)	0.95 (0.77 - 1.18)	0.95 (0.77 - 1.18)
	3 (26.17-30.47)	372	366	0.96 (0.77 - 1.18)	0.95 (0.76 - 1.19)	0.98 (0.79 - 1.23)
	4 > 30.47	391	427	0.85 (0.69 - 1.05)	0.85 (0.69 - 1.06)	0.85 (0.69 - 1.06)
	P for trend[e]			0.12	0.09	0.14
Former smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (5.08-29.89)	164	158	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (29.90-49.90)	161	165	0.94 (0.69 - 1.28)	0.95 (0.68 - 1.31)	0.95 (0.68 - 1.32)
	3 (49.91-90.05)	178	181	0.95 (0.69 - 1.29)	0.94 (0.67 - 1.3)	0.9 (0.65 - 1.26)
	4 > 90.05	200	199	0.97 (0.71 - 1.32)	1.03 (0.73 - 1.43)	1.08 (0.77 - 1.51)
	P for trend[e]			0.32	0.52	0.61
	Folate (Vitamin B9) [nmol/L][d]					
	1 (1.70-17.45)	110	100	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (17.46-33.35)	163	159	0.92 (0.63 - 1.32)	0.94 (0.64 - 1.38)	1.02 (0.7 - 1.51)
	3 (33.36-59.10)	196	210	0.84 (0.59 - 1.19)	0.87 (0.6 - 1.26)	0.94 (0.65 - 1.37)
	4 > 59.10	234	234	0.89 (0.62 - 1.28)	0.97 (0.66 - 1.43)	1.02 (0.7 - 1.5)
	P for trend[e]			0.51	0.83	0.88
	Methionine [µmol/L][d]					
	1 (11.83-22.77)	150	132	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.78-26.16)	165	167	0.87 (0.63 - 1.2)	0.92 (0.65 - 1.3)	0.89 (0.63 - 1.27)
	3 (26.17-30.47)	184	189	0.85 (0.62 - 1.16)	0.85 (0.6 - 1.18)	0.93 (0.66 - 1.32)
	4 > 30.47	204	215	0.83 (0.61 - 1.13)	0.87 (0.62 - 1.21)	0.88 (0.63 - 1.24)
	P for trend[e]			0.13	0.16	0.28
Current smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (5.08-29.89)	323	297	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (29.90-49.90)	225	222	0.98 (0.75 - 1.28)	1 (0.77 - 1.3)	0.98 (0.75 - 1.28)
	3 (49.91-90.05)	150	162	0.91 (0.69 - 1.21)	0.94 (0.7 - 1.24)	0.91 (0.69 - 1.21)
	4 > 90.05	122	139	0.85 (0.63 - 1.15)	0.87 (0.65 - 1.18)	0.86 (0.64 - 1.17)
	P for trend[e]			0.52	0.68	0.57
	Folate (Vitamin B9) [nmol/L][d]					
	1 (1.70-17.45)	340	329	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (17.46-33.35)	206	229	0.91 (0.71 - 1.17)	0.93 (0.73 - 1.2)	0.91 (0.71 - 1.17)
	3 (33.36-59.10)	157	140	1.07 (0.79 - 1.44)	1.09 (0.8 - 1.47)	1.06 (0.78 - 1.44)
	4 > 59.10	117	122	1.06 (0.75 - 1.51)	1.06 (0.74 - 1.51)	1.06 (0.75 - 1.52)
	P for trend[e]	.	.	0.97	0.95	0.98
	Methionine [µmol/L][d]					
	1 (11.83-22.77)	235	231	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.78-26.16)	210	200	1.01 (0.77 - 1.33)	1 (0.76 - 1.32)	1.01 (0.76 - 1.33)
	3 (26.17-30.47)	188	177	1.03 (0.76 - 1.38)	1.03 (0.76 - 1.38)	1.03 (0.76 - 1.38)
	4 > 30.47	187	212	0.85 (0.63 - 1.14)	0.84 (0.62 - 1.12)	0.84 (0.62 - 1.12)
	P for trend[e]			0.39	0.32	0.36

	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]					
	1 (5.08-29.89)	221	192	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (29.90-49.90)	175	174	0.89 (0.66 - 1.2)	0.91 (0.67 - 1.24)	0.87 (0.64 - 1.18)
	3 (49.91-90.05)	150	156	0.83 (0.6 - 1.13)	0.82 (0.59 - 1.14)	0.74 (0.53 - 1.03)
	4 > 90.05	129	153	0.7 (0.5 - 0.98)	0.74 (0.53 - 1.06)	0.75 (0.52 - 1.06)
	P for trend[e]			0.02	0.04	0.04
	Folate (Vitamin B9) [nmol/L][d]					
	1 (1.70-17.45)	185	164	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (17.46-33.35)	152	176	0.8 (0.57 - 1.11)	0.85 (0.6 - 1.18)	0.85 (0.6 - 1.19)
Smoking men	3 (33.36-59.10)	172	166	0.91 (0.64 - 1.32)	0.98 (0.67 - 1.42)	0.98 (0.67 - 1.43)
	4 > 59.10	166	169	0.89 (0.6 - 1.32)	0.96 (0.64 - 1.45)	0.97 (0.64 - 1.47)
	P for trend[e]			0.93	0.80	0.87
	Methionine [µmol/L][d]					
	1 (11.83-22.77)	161	147	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.78-26.16)	161	133	1.06 (0.76 - 1.49)	1.05 (0.75 - 1.49)	1.03 (0.73 - 1.47)
	3 (26.17-30.47)	162	155	0.96 (0.69 - 1.33)	0.94 (0.67 - 1.33)	0.99 (0.69 - 1.4)
	4 > 30.47	191	240	0.69 (0.5 - 0.94)	0.69 (0.49 - 0.95)	0.68 (0.49 - 0.95)
	P for trend[e]			0.02	0.01	0.03
	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]					
	1 (5.08-29.89)	266	263	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (29.90-49.90)	211	213	1.02 (0.77 - 1.33)	1.03 (0.78 - 1.36)	1.03 (0.78 - 1.36)
	3 (49.91-90.05)	178	187	1 (0.76 - 1.32)	1.03 (0.78 - 1.37)	1.05 (0.79 - 1.4)
	4 > 90.05	193	185	1.12 (0.84 - 1.49)	1.15 (0.86 - 1.54)	1.16 (0.86 - 1.55)
	P for trend[e]			0.52	0.34	0.35
	Folate (Vitamin B9) [nmol/L][d]					
	1 (1.70-17.45)	265	265	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (17.46-33.35)	217	212	1.05 (0.8 - 1.36)	1.07 (0.82 - 1.39)	1.04 (0.8 - 1.36)
Smoking women	3 (33.36-59.10)	181	184	1.01 (0.76 - 1.34)	1.03 (0.77 - 1.37)	1.01 (0.76 - 1.34)
	4 > 59.10	185	187	1.05 (0.77 - 1.42)	1.08 (0.79 - 1.48)	1.07 (0.79 - 1.46)
	P for trend[e]			0.77	0.92	0.91
	Methionine [µmol/L][d]					
	1 (11.83-22.77)	224	216	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.78-26.16)	214	234	0.86 (0.65 - 1.13)	0.88 (0.66 - 1.16)	0.89 (0.68 - 1.18)
	3 (26.17-30.47)	210	211	0.96 (0.73 - 1.28)	0.97 (0.72 - 1.3)	1 (0.75 - 1.33)
	4 > 30.47	200	187	1.06 (0.79 - 1.41)	1.07 (0.79 - 1.43)	1.07 (0.8 - 1.44)
	P for trend			0.85	0.86	0.81

a) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles) and education (in 7 categories).

b) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles), education (in 7 categories), and number of years of smoking (continuous).

c) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles), education (in 7 categories), and number of pack-years of smoking (continuous).

d) Quartile cut-off points were determined based on the circulating level distribution of US matched controls.

e) P for trend assessed by the base 2 logarithm of the circulating levels.

Supplementary Table 2. Odds ratios of lung cancer for circulating levels of vitamin B6, folate and methionine among ever smoking participants in EU/AU

		European/Asian cohorts				
				Model 1[a]	Model 2[b]	Model 3[c]
		Cases	Controls	OR (95%CI)	OR (95%CI)	OR (95%CI)
All ever smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (4.37-22.41)	334	245	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.42-30.90)	216	254	0.65 (0.5 - 0.83)	0.63 (0.49 - 0.82)	0.65 (0.5 - 0.84)
	3 (30.91-44.95)	186	217	0.65 (0.5 - 0.86)	0.65 (0.49 - 0.86)	0.66 (0.5 - 0.87)
	4 > 44.96	201	221	0.71 (0.54 - 0.93)	0.71 (0.54 - 0.94)	0.69 (0.52 - 0.91)
	P for trend[e]			0.31	0.65	0.17
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.24-7.78)	279	230	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (7.79-10.77)	220	243	0.74 (0.56 - 0.97)	0.77 (0.58 - 1.01)	0.74 (0.56 - 0.97)
	3 (10.78-15.98)	230	231	0.84 (0.64 - 1.11)	0.83 (0.63 - 1.1)	0.86 (0.65 - 1.13)
	4 > 15.98	208	233	0.76 (0.57 - 1)	0.76 (0.57 - 1.02)	0.75 (0.56 - 1)
	P for trend[e]			0.06	0.06	0.04
	Methionine [µmol/L][d]					
	1 (10.63-22.81)	221	235	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.82-26.29)	249	218	1.23 (0.94 - 1.6)	1.18 (0.89 - 1.55)	1.24 (0.95 - 1.63)
	3 (26.30-29.91)	235	241	0.97 (0.74 - 1.29)	1 (0.75 - 1.33)	0.99 (0.75 - 1.31)
	4 > 29.91	232	243	1.01 (0.77 - 1.33)	0.97 (0.73 - 1.3)	1.03 (0.78 - 1.37)
	P for trend[e]			0.79	0.75	0.96
Former smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (4.37-22.41)	48	35	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.42-30.90)	62	55	0.8 (0.45 - 1.42)	0.88 (0.47 - 1.63)	0.83 (0.45 - 1.52)
	3 (30.91-44.95)	58	74	0.55 (0.31 - 0.98)	0.64 (0.34 - 1.2)	0.54 (0.29 - 1)
	4 > 44.96	77	81	0.66 (0.38 - 1.16)	0.67 (0.36 - 1.23)	0.61 (0.34 - 1.09)
	P for trend[e]			0.72	0.51	0.28
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.24-7.78)	50	31	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (7.79-10.77)	53	60	0.52 (0.28 - 0.95)	0.53 (0.27 - 1.02)	0.55 (0.29 - 1.03)
	3 (10.78-15.98)	62	64	0.57 (0.32 - 1.02)	0.54 (0.29 - 1.02)	0.58 (0.31 - 1.06)
	4 > 15.98	80	90	0.51 (0.28 - 0.91)	0.48 (0.25 - 0.9)	0.5 (0.27 - 0.92)
	P for trend[e]			0.03	0.02	0.01
	Methionine [µmol/L][d]					
	1 (10.63-22.81)	52	64	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.82-26.29)	56	51	1.33 (0.79 - 2.24)	1.35 (0.76 - 2.39)	1.31 (0.76 - 2.24)
	3 (26.30-29.91)	68	61	1.37 (0.82 - 2.29)	1.54 (0.88 - 2.69)	1.28 (0.75 - 2.19)
	4 > 29.91	69	69	1.23 (0.75 - 2.03)	1.05 (0.61 - 1.82)	1.21 (0.72 - 2.05)
	P for trend[e]			0.82	0.91	0.93
Current smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (4.37-22.41)	286	210	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.42-30.90)	154	199	0.6 (0.45 - 0.8)	0.58 (0.43 - 0.77)	0.6 (0.45 - 0.8)
	3 (30.91-44.95)	128	143	0.7 (0.51 - 0.96)	0.68 (0.49 - 0.94)	0.7 (0.51 - 0.97)
	4 > 44.96	124	140	0.74 (0.54 - 1.02)	0.76 (0.55 - 1.05)	0.74 (0.54 - 1.02)
	P for trend[e]			0.33	0.44	0.31
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.24-7.78)	229	199	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (7.79-10.77)	167	183	0.79 (0.59 - 1.08)	0.82 (0.6 - 1.11)	0.78 (0.58 - 1.07)
	3 (10.78-15.98)	168	167	0.91 (0.67 - 1.26)	0.91 (0.66 - 1.25)	0.92 (0.67 - 1.27)
	4 > 15.98	128	143	0.83 (0.59 - 1.16)	0.84 (0.6 - 1.18)	0.82 (0.59 - 1.15)
	P for trend[e]			0.39	0.43	0.36
	Methionine [µmol/L][d]					
	1 (10.63-22.81)	169	171	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.82-26.29)	193	167	1.17 (0.85 - 1.6)	1.13 (0.82 - 1.55)	1.19 (0.87 - 1.64)
	3 (26.30-29.91)	167	180	0.84 (0.61 - 1.17)	0.85 (0.61 - 1.19)	0.88 (0.63 - 1.23)
	4 > 29.91	163	174	0.92 (0.66 - 1.28)	0.92 (0.66 - 1.29)	0.95 (0.68 - 1.33)
	P for trend[e]			0.60	0.66	0.81

	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]					
	1 (4.37-22.41)	195	146	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.42-30.90)	140	162	0.67 (0.49 - 0.92)	0.68 (0.49 - 0.94)	0.7 (0.5 - 0.97)
	3 (30.91-44.95)	123	133	0.73 (0.51 - 1.03)	0.76 (0.53 - 1.09)	0.74 (0.51 - 1.06)
	4 > 44.96	117	134	0.72 (0.51 - 1.02)	0.74 (0.52 - 1.06)	0.71 (0.5 - 1.02)
	P for trend[e]			0.29	0.37	0.21
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.24-7.78)	158	121	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (7.79-10.77)	135	145	0.71 (0.5 - 1.02)	0.72 (0.5 - 1.04)	0.71 (0.49 - 1.02)
Smoking men	3 (10.78-15.98)	151	161	0.73 (0.51 - 1.04)	0.69 (0.48 - 1)	0.75 (0.52 - 1.08)
	4 > 15.98	131	148	0.67 (0.46 - 0.97)	0.66 (0.45 - 0.98)	0.65 (0.44 - 0.96)
	P for trend[e]			0.03	0.02	0.01
	Methionine [µmol/L][d]					
	1 (10.63-22.81)	95	100	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.82-26.29)	144	126	1.23 (0.84 - 1.8)	1.22 (0.82 - 1.81)	1.29 (0.87 - 1.91)
	3 (26.30-29.91)	164	166	0.98 (0.67 - 1.44)	1.06 (0.71 - 1.58)	1.03 (0.7 - 1.52)
	4 > 29.91	172	183	0.99 (0.68 - 1.42)	1 (0.68 - 1.46)	1.03 (0.7 - 1.5)
	P for trend[e]			0.53	0.65	0.73
	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]					
	1 (4.37-22.41)	139	99	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.42-30.90)	76	92	0.59 (0.39 - 0.9)	0.54 (0.35 - 0.83)	0.55 (0.35 - 0.84)
	3 (30.91-44.95)	63	84	0.54 (0.35 - 0.85)	0.5 (0.31 - 0.79)	0.54 (0.34 - 0.85)
	4 > 44.96	84	87	0.67 (0.43 - 1.04)	0.66 (0.42 - 1.03)	0.64 (0.41 - 0.99)
	P for trend[e]			0.69	0.67	0.46
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.24-7.78)	121	109	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (7.79-10.77)	85	98	0.76 (0.5 - 1.16)	0.82 (0.53 - 1.25)	0.78 (0.51 - 1.19)
Smoking women	3 (10.78-15.98)	79	70	1.06 (0.67 - 1.68)	1.12 (0.7 - 1.79)	1.06 (0.66 - 1.69)
	4 > 15.98	77	85	0.88 (0.57 - 1.37)	0.9 (0.57 - 1.4)	0.9 (0.57 - 1.4)
	P for trend[e]			0.81	0.89	0.88
	Methionine [µmol/L][d]					
	1 (10.63-22.81)	126	135	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.82-26.29)	105	92	1.21 (0.82 - 1.78)	1.15 (0.77 - 1.71)	1.19 (0.8 - 1.76)
	3 (26.30-29.91)	71	75	0.95 (0.62 - 1.45)	0.93 (0.6 - 1.44)	0.94 (0.61 - 1.45)
	4 > 29.91	60	60	1.06 (0.67 - 1.67)	0.96 (0.61 - 1.54)	1.07 (0.67 - 1.7)
	P for trend[e]			0.83	0.99	0.80

a) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles) and education (in 7 categories).

b) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles), education (in 7 categories), and number of years of smoking (continuous).

c) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles), education (in 7 categories), and number of pack-years of smoking (continuous).

d) Quartile cut-off points were determined based on the circulating level distribution of EU/AU matched controls.

e) P for trend assessed by the base 2 logarithm of the circulating levels.

Supplementary Table 3. Odds ratios of lung cancer for circulating levels of vitamin B6, folate and methionine among ever smoking participants in Asia

		Asian cohorts				
				Model 1[a]	**Model 2[b]**	**Model 3[c]**
		Cases	Controls	OR (95%CI)	OR (95%CI)	OR (95%CI)
All ever smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (4.81-20.49)	490	391	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (20.50-30.82)	317	308	0.89 (0.71 - 1.13)	0.91 (0.72 - 1.15)	0.96 (0.76 - 1.22)
	3 (31.23-46.99)	205	259	0.68 (0.52 - 0.88)	0.72 (0.55 - 0.94)	0.73 (0.55 - 0.95)
	4 > 47.00	161	215	0.65 (0.49 - 0.87)	0.71 (0.54 - 0.95)	0.73 (0.54 - 0.97)
	P for trend[e]			0.001	0.009	0.01
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.59)	418	348	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.60-15.00)	315	310	0.85 (0.68 - 1.07)	0.89 (0.71 - 1.13)	0.91 (0.72 - 1.15)
	3 (15.01-21.00)	222	274	0.73 (0.58 - 0.94)	0.77 (0.6 - 0.99)	0.81 (0.62 - 1.04)
	4 >21.01	218	241	0.83 (0.64 - 1.06)	0.89 (0.69 - 1.15)	0.89 (0.68 - 1.15)
	P for trend[e]	.	.	0.05	0.26	0.22
	Methionine [µmol/L][d]					
	1 (12.32-23.17)	258	254	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.18-27.39)	282	267	1.12 (0.86 - 1.45)	1.22 (0.93 - 1.59)	1.19 (0.91 - 1.56)
	3 (27.39-32.22)	288	312	0.93 (0.71 - 1.2)	0.98 (0.75 - 1.29)	0.94 (0.72 - 1.23)
	4 >32.22	345	340	1.03 (0.8 - 1.31)	1.08 (0.83 - 1.39)	1.08 (0.83 - 1.39)
	P for trend[e]			0.87	0.99	0.83
Former smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (4.81-20.49)	34	22	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (20.50-30.82)	40	38	0.67 (0.32 - 1.41)	0.71 (0.33 - 1.51)	0.82 (0.36 - 1.83)
	3 (31.23-46.99)	48	52	0.6 (0.3 - 1.17)	0.6 (0.3 - 1.2)	0.63 (0.31 - 1.31)
	4 > 47.00	54	64	0.55 (0.29 - 1.06)	0.64 (0.33 - 1.26)	0.66 (0.33 - 1.32)
	P for trend[e]			0.29	0.58	0.65
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.59)	60	37	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.60-15.00)	32	48	0.42 (0.23 - 0.78)	0.47 (0.25 - 0.88)	0.47 (0.25 - 0.91)
	3 (15.01-21.00)	34	45	0.45 (0.24 - 0.86)	0.53 (0.27 - 1.02)	0.61 (0.31 - 1.2)
	4 >21.01	50	46	0.65 (0.36 - 1.19)	0.82 (0.43 - 1.55)	0.85 (0.45 - 1.62)
	P for trend[e]			0.26	0.74	0.76
	Methionine [µmol/L][d]					
	1 (12.32-23.17)	62	48	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.18-27.39)	37	41	0.66 (0.36 - 1.22)	0.74 (0.4 - 1.4)	0.77 (0.4 - 1.47)
	3 (27.39-32.22)	43	45	0.69 (0.38 - 1.25)	0.74 (0.4 - 1.36)	0.71 (0.38 - 1.32)
	4 >32.22	34	42	0.59 (0.32 - 1.1)	0.61 (0.32 - 1.15)	0.56 (0.28 - 1.09)
	P for trend[e]			0.17	0.17	0.17
Current smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (4.81-20.49)	456	369	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (20.50-30.82)	277	270	0.93 (0.73 - 1.18)	0.93 (0.73 - 1.19)	0.98 (0.76 - 1.26)
	3 (31.23-46.99)	157	207	0.68 (0.51 - 0.9)	0.74 (0.55 - 0.99)	0.73 (0.55 - 0.99)
	4 > 47.00	107	151	0.68 (0.49 - 0.94)	0.7 (0.5 - 0.98)	0.73 (0.52 - 1.02)
	P for trend[e]			0.001	0.006	0.009
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.59)	358	311	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.60-15.00)	283	262	0.96 (0.75 - 1.23)	0.98 (0.76 - 1.26)	1 (0.78 - 1.29)
	3 (15.01-21.00)	188	229	0.8 (0.61 - 1.05)	0.82 (0.62 - 1.07)	0.84 (0.64 - 1.11)
	4 >21.01	168	195	0.84 (0.64 - 1.11)	0.85 (0.64 - 1.13)	0.86 (0.65 - 1.15)
	P for trend[e]			0.09	0.18	0.18
	Methionine [µmol/L][d]					
	1 (12.32-23.17)	196	206	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.18-27.39)	245	226	1.28 (0.96 - 1.72)	1.37 (1.02 - 1.86)	1.33 (0.98 - 1.8)
	3 (27.39-32.22)	245	267	1.02 (0.76 - 1.36)	1.07 (0.79 - 1.45)	1.02 (0.75 - 1.38)
	4 >32.22	311	298	1.16 (0.88 - 1.52)	1.21 (0.91 - 1.6)	1.22 (0.92 - 1.61)
	P for trend[e]			0.71	0.58	0.44

	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]					
	1 (4.81-20.49)	473	380	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (20.50-30.82)	293	292	0.89 (0.7 - 1.12)	0.89 (0.7 - 1.13)	0.96 (0.75 - 1.22)
	3 (31.23-46.99)	183	233	0.69 (0.53 - 0.9)	0.72 (0.55 - 0.95)	0.74 (0.56 - 0.98)
	4 > 47.00	151	195	0.69 (0.51 - 0.92)	0.75 (0.56 - 1.01)	0.77 (0.56 - 1.04)
	P for trend[e]			0.003	0.02	0.04
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.59)	400	328	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.60-15.00)	296	292	0.84 (0.66 - 1.05)	0.88 (0.69 - 1.12)	0.89 (0.7 - 1.13)
Smoking men	3 (15.01-21.00)	208	250	0.75 (0.58 - 0.96)	0.8 (0.62 - 1.04)	0.82 (0.63 - 1.07)
	4 >21.01	196	230	0.76 (0.59 - 0.98)	0.82 (0.63 - 1.07)	0.81 (0.62 - 1.06)
	P for trend[e]			0.02	0.15	0.09
	Methionine [µmol/L][d]					
	1 (12.32-23.17)	225	225	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.18-27.39)	261	250	1.11 (0.85 - 1.46)	1.22 (0.92 - 1.61)	1.19 (0.89 - 1.58)
	3 (27.39-32.22)	279	301	0.94 (0.72 - 1.24)	1.01 (0.76 - 1.33)	0.96 (0.72 - 1.27)
	4 >32.22	335	324	1.06 (0.82 - 1.38)	1.12 (0.86 - 1.46)	1.12 (0.86 - 1.47)
	P for trend[e]			0.97	0.88	0.67
	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]					
	1 (4.81-20.49)	17	11	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (20.50-30.82)	24	16	0.71 (0.22 - 2.26)	0.82 (0.25 - 2.7)	0.78 (0.24 - 2.51)
	3 (31.23-46.99)	22	26	0.37 (0.12 - 1.17)	0.47 (0.14 - 1.57)	0.4 (0.12 - 1.3)
	4 > 47.00	10	20	0.29 (0.08 - 1.04)	0.32 (0.09 - 1.18)	0.31 (0.08 - 1.14)
	P for trend[e]			0.04	0.07	0.05
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.59)	18	20	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.60-15.00)	19	18	1.3 (0.39 - 4.33)	1.63 (0.46 - 5.79)	1.66 (0.46 - 5.99)
Smoking women	3 (15.01-21.00)	14	24	0.65 (0.22 - 1.91)	0.59 (0.19 - 1.84)	0.78 (0.25 - 2.41)
	4 >21.01	22	11	3 (0.9 - 9.93)	3.96 (1.03 - 15.18)	3.86 (1.05 - 14.21)
	P for trend[e]			0.17	0.16	0.11
	Methionine [µmol/L][d]					
	1 (12.32-23.17)	33	29	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.18-27.39)	21	17	1.21 (0.44 - 3.34)	1.35 (0.48 - 3.79)	1.27 (0.45 - 3.56)
	3 (27.39-32.22)	9	11	0.64 (0.2 - 2.05)	0.7 (0.22 - 2.24)	0.75 (0.23 - 2.42)
	4 >32.22	10	16	0.51 (0.2 - 1.33)	0.54 (0.2 - 1.44)	0.54 (0.2 - 1.44)
	P for trend[e]			0.34	0.42	0.38

a) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles) and education (in 7 categories).

b) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles), education (in 7 categories), and number of years of smoking (continuous).

c) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles), education (in 7 categories), and number of pack-years of smoking (continuous).

d) Quartile cut-off points were determined based on the circulating level distribution of Asian matched controls.

e) P for trend assessed by the base 2 logarithm of the circulating levels.

Supplementary Table 4. Odds ratios of lung cancer for circulating levels of vitamin B6, folate and methionine among ever smoking participants in LC3

		All cohorts				
				Model 1[a]	Model 2[b]	Model 3[c]
		Cases	Controls	OR (95%CI)	OR (95%CI)	OR (95%CI)
All ever smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (4.37-23.93)	1300	1104	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.94-37.05)	911	962	0.83 (0.73 - 0.94)	0.85 (0.74 - 0.97)	0.87 (0.76 - 0.99)
	3 (37.06-62.50)	724	795	0.81 (0.7 - 0.93)	0.83 (0.72 - 0.96)	0.82 (0.71 - 0.95)
	4 > 62.51	698	772	0.78 (0.67 - 0.91)	0.83 (0.71 - 0.97)	0.81 (0.7 - 0.95)
	P for trend[e]			0.003	0.03	0.02
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.92)	1154	1039	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.93-17.89)	948	957	0.91 (0.8 - 1.04)	0.93 (0.81 - 1.06)	0.94 (0.82 - 1.08)
	3 (17.90-34.92)	776	855	0.85 (0.74 - 0.98)	0.88 (0.76 - 1.02)	0.89 (0.77 - 1.03)
	4 > 34.97	755	782	0.87 (0.73 - 1.03)	0.92 (0.77 - 1.1)	0.91 (0.76 - 1.08)
	P for trend[e]			0.03	0.15	0.10
	Methionine [µmol/L][d]					
	1 (10.63-22.89)	861	855	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.90-27.40)	902	849	1.07 (0.93 - 1.23)	1.09 (0.94 - 1.25)	1.09 (0.94 - 1.25)
	3 (27.40-31.00)	913	913	0.99 (0.86 - 1.14)	1.01 (0.87 - 1.17)	1.02 (0.88 - 1.18)
	4 > 31.01	957	1016	0.95 (0.82 - 1.09)	0.96 (0.83 - 1.11)	0.96 (0.83 - 1.11)
	P for trend[e]			0.25	0.24	0.44
Former smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (4.37-23.93)	210	163	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.94-37.05)	266	279	0.74 (0.57 - 0.97)	0.77 (0.58 - 1.01)	0.77 (0.58 - 1.03)
	3 (37.06-62.50)	270	288	0.72 (0.55 - 0.94)	0.76 (0.57 - 1.01)	0.7 (0.53 - 0.94)
	4 > 62.51	378	394	0.73 (0.56 - 0.95)	0.79 (0.6 - 1.04)	0.77 (0.58 - 1.02)
	P for trend[e]			0.17	0.41	0.31
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.92)	211	164	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.93-17.89)	200	215	0.69 (0.52 - 0.93)	0.74 (0.54 - 1.01)	0.74 (0.54 - 1)
	3 (17.90-34.92)	275	281	0.69 (0.51 - 0.92)	0.77 (0.56 - 1.05)	0.79 (0.57 - 1.08)
	4 > 34.97	438	464	0.63 (0.47 - 0.86)	0.73 (0.53 - 1.01)	0.72 (0.52 - 1)
	P for trend[e]			0.05	0.26	0.18
	Methionine [µmol/L][d]					
	1 (10.63-22.89)	263	248	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.90-27.40)	279	267	0.98 (0.77 - 1.26)	1.01 (0.78 - 1.31)	1.01 (0.77 - 1.31)
	3 (27.40-31.00)	298	300	0.93 (0.73 - 1.18)	0.96 (0.75 - 1.24)	0.99 (0.76 - 1.28)
	4 > 31.01	284	309	0.86 (0.67 - 1.1)	0.86 (0.66 - 1.12)	0.87 (0.67 - 1.13)
	P for trend[e]			0.09	0.08	0.15
Current smokers	**Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]**					
	1 (4.37-23.93)	1090	941	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.94-37.05)	645	683	0.85 (0.74 - 0.99)	0.87 (0.75 - 1.01)	0.89 (0.76 - 1.03)
	3 (37.06-62.50)	454	507	0.83 (0.7 - 0.98)	0.85 (0.72 - 1.01)	0.85 (0.72 - 1.01)
	4 > 62.51	320	378	0.78 (0.64 - 0.95)	0.81 (0.67 - 0.99)	0.8 (0.65 - 0.97)
	P for trend[e]			0.005	0.03	0.01
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.92)	943	875	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.93-17.89)	748	742	0.96 (0.83 - 1.12)	0.97 (0.83 - 1.13)	0.98 (0.84 - 1.14)
	3 (17.90-34.92)	501	574	0.88 (0.74 - 1.04)	0.89 (0.75 - 1.05)	0.9 (0.76 - 1.06)
	4 > 34.97	317	318	0.97 (0.78 - 1.21)	1.01 (0.8 - 1.26)	0.98 (0.78 - 1.23)
	P for trend[e]			0.15	0.26	0.20
	Methionine [µmol/L][d]					
	1 (10.63-22.89)	598	607	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.90-27.40)	623	582	1.12 (0.94 - 1.32)	1.12 (0.95 - 1.33)	1.12 (0.95 - 1.33)
	3 (27.40-31.00)	615	613	1.02 (0.86 - 1.21)	1.03 (0.86 - 1.22)	1.03 (0.87 - 1.23)
	4 > 31.01	673	707	0.99 (0.84 - 1.17)	1 (0.84 - 1.19)	1 (0.84 - 1.19)
	P for trend[e]			0.72	0.72	0.92

	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]					
	1 (4.37-23.93)	952	789	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.94-37.05)	593	655	0.76 (0.65 - 0.89)	0.79 (0.67 - 0.92)	0.82 (0.7 - 0.96)
	3 (37.06-62.50)	461	497	0.81 (0.68 - 0.96)	0.84 (0.7 - 1.01)	0.82 (0.69 - 0.99)
	4 > 62.51	344	409	0.68 (0.56 - 0.84)	0.72 (0.59 - 0.89)	0.7 (0.57 - 0.87)
	P for trend[e]			<.0001	0.001	0.001
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.92)	799	693	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.93-17.89)	677	687	0.88 (0.75 - 1.03)	0.9 (0.77 - 1.06)	0.92 (0.78 - 1.08)
Smoking men	3 (17.90-34.92)	489	562	0.78 (0.66 - 0.94)	0.82 (0.68 - 0.99)	0.84 (0.7 - 1.01)
	4 > 34.97	385	408	0.78 (0.61 - 0.98)	0.86 (0.67 - 1.1)	0.82 (0.64 - 1.05)
	P for trend[e]			0.005	0.05	0.02
	Methionine [µmol/L][d]					
	1 (10.63-22.89)	476	464	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.90-27.40)	538	490	1.11 (0.92 - 1.33)	1.14 (0.94 - 1.38)	1.13 (0.93 - 1.37)
	3 (27.40-31.00)	618	624	0.97 (0.81 - 1.17)	1.02 (0.84 - 1.23)	1 (0.83 - 1.21)
	4 > 31.01	718	772	0.92 (0.77 - 1.1)	0.94 (0.79 - 1.13)	0.94 (0.78 - 1.13)
	P for trend[e]			0.16	0.17	0.32
	Vitamin B6 (Pyridoxal 5'-phosphate) [nmol/L][d]					
	1 (4.37-23.93)	348	315	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (23.94-37.05)	318	307	0.96 (0.77 - 1.2)	0.98 (0.78 - 1.23)	0.96 (0.76 - 1.2)
	3 (37.06-62.50)	263	298	0.81 (0.64 - 1.03)	0.83 (0.65 - 1.06)	0.82 (0.65 - 1.05)
	4 > 62.51	354	363	0.94 (0.74 - 1.19)	0.98 (0.77 - 1.25)	0.97 (0.77 - 1.24)
	P for trend[e]			0.96	0.71	0.91
	Folate (Vitamin B9) [nmol/L][d]					
	1 (0.17-10.92)	355	346	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (10.93-17.89)	271	270	0.98 (0.77 - 1.25)	0.98 (0.76 - 1.25)	0.97 (0.76 - 1.24)
Smoking women	3 (17.90-34.92)	287	293	0.99 (0.78 - 1.28)	1 (0.77 - 1.28)	1 (0.77 - 1.28)
	4 > 34.97	370	374	1.01 (0.78 - 1.31)	1.03 (0.79 - 1.34)	1.03 (0.79 - 1.33)
	P for trend[e]			0.89	0.93	0.90
	Methionine [µmol/L][d]					
	1 (10.63-22.89)	385	391	1.00 (reference)	1.00 (reference)	1.00 (reference)
	2 (22.90-27.40)	364	359	1.02 (0.83 - 1.26)	1.02 (0.82 - 1.26)	1.03 (0.83 - 1.28)
	3 (27.40-31.00)	295	289	1.03 (0.82 - 1.29)	1.01 (0.8 - 1.27)	1.05 (0.83 - 1.33)
	4 > 31.01	239	244	1.02 (0.8 - 1.29)	1.02 (0.8 - 1.31)	1.02 (0.8 - 1.3)
	P for trend[e]			0.97	0.97	0.93

a) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles) and education (in 7 categories).

b) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles), education (in 7 categories), and number of years of smoking (continuous).

c) Assessed by conditional logistic regression, conditioning on individual case set and adjusting for circulating cotinine (in quartiles), education (in 7 categories), and number of pack-years of smoking (continuous).

d) Quartile cut-off points were determined based on the circulating level distribution of LC3 matched controls.

e) P for trend assessed by the base 2 logarithm of the circulating levels.

Supplementary table 5. Circulating concentrations of vitamin B6, folate and methionine among users vs non-users of multivitamins

Multivitamins users	Nb.	Analyte	Median	5th	95th
No	2703	Vitamin B6	39.21	14.35	186.5
		Folate	25.51	7.619	104.56
		Methionine	26.05	18.32	40.07
Yes	1607	Vitamin B6	76.18	22.43	344.1
		Folate	46.52	10.741	125.1
		Methionine	25.79	18.903	38.86

Supplementary table 6. Intraclass Correlation Coefficient for vitamin B6, folate and methionine as measured using repeat samples from CPS-II and NSHDS cohorts.

LC3 Participants	Time between blood draws	Intraclass correlation coefficient		
		Vitamin B6	Folate	Methionine
CSP-II and NHSDS cohorts	6 months	0.73	0.58	0.40
	2-6 years	0.62	0.34	0.27
	10 years	0.33	0.49	0.51

Supplementary Methods

Material and Methods

Study population

We invited all prospective cohort studies that in 2009 were members in the US National Cancer Institute (NCI) Cohort Consortium to participate in the study. Additional inclusion criteria included the occurrence of at least 200 incident lung cancer cases with baseline questionnaire data and either plasma or serum samples cryopreserved at <80°C available. Twenty cohorts fulfilled those criteria and accepted to participate, resulting in a combined cohort population of over 2,000,000 participants from North America, Europe, Asia and Australia.

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 5,545 lung cancer cases, and in order optimize the statistical power in smoking stratified analyses, 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), 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 years, ≥15 cigarettes per day). After excluding cases who were not correctly matched on smoking status (n=124 cases), who had insufficient plasma samples (n=42), or had a revised date of diagnosis prior to blood draw (n=13), 5,364 lung cancer case-control pairs remained eligible for the risk analysis.

Brief description of the participating cohorts in the Lung Cancer Cohort Consortium

1) **US cohorts**

The Women's Health Initiative (WHI)

WHI is a long-term health study of 161,808 post-menopausal women aged 50 to 79 years at 40 clinical centers throughout the U.S. WHI comprises a Clinical Trial (CT) component (68,132 women), and an Observational Study (OS) component (93,676 women), and has included several extension studies. Some detailed descriptions of WHI have been previously presented[1] [2].The CT evaluated two forms of postmenopausal hormone therapy, a low-fat dietary pattern intervention, and calcium and vitamin D supplementation in a randomized, controlled fashion, in a partial factorial design. The hormone therapy component findings led to major reductions in the use of hormone therapy worldwide, and are thought to have led to noteworthy reductions in breast cancer incidence.

In the present study lung cancer cases occurring during the follow-up of WHI cohorts since enrolment (1993-1998) among non-smoking women, were matched 1-1 to corresponding non-smoking lung cancer free controls, for serum and DNA analyte comparisons.

The Southern Community Cohort Study (SCCS)

The Southern Community Cohort Study (SCCS)[3] is a prospective cohort of African and non-African Americans which during 2002-2009 enrolled approximately 86,000 residents aged 40-79 years across 12 southern states. Recruitment occurred mainly at community health centers, institutions providing basic health services primarily to the medically uninsured, so that the cohort includes many adults of lower income and educational status. Each study participant completed a detailed baseline questionnaire, and nearly 90% provided a biologic specimen (approximately 45% a blood sample and 45% buccal cells). Follow-up of the cohort is conducted by linkage to national mortality registers and to state cancer registries. Included in this study are 240 incident African American lung cancer cases and 240 individually matched African American cohort members without lung cancer at the index date selected by incidence density sampling.

Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO)

The PLCO study, a randomized trial aimed at evaluating the efficacy of screening in reducing cancer mortality, recruited approximately 155,000 men and women age 55 to 74 years from 1992 to 2001[4]. Screening for lung cancer among participants in the intervention arm included a chest x-ray at baseline followed by either three annual x-rays (for current or

former smokers at enrollment) or two annual x-rays (for never smokers); participants in the control arm received routine health care. Screening-arm participants provided data on sociodemographic factors, smoking behavior, anthropometric characteristics, medical history, and family history of cancer, as well as blood samples annually for the first 6 years of the study (baseline [T0] and T1 through T5). Lung cancers were ascertained through annual questionnaires mailed to the participants, and positive reports were followed up by abstracting medical records or death certificates. Follow-up in the trial as of July 2009 was 96.7%.

We conducted a nested case-control study within the screening arm of the PLCO trial. As of December 31, 2004, 898 lung cancers were diagnosed among the 77,464 participants. Patients were excluded because of missing baseline questionnaire, previous history of any cancer, diagnosis of multiple cancers during follow-up, missing smoking information at baseline, missing consent for utilization of biologic specimens for etiologic studies, or unavailability/insufficient quantity of serum or DNA specimens. Hemolyzed vials were excluded. We included 450 confirmed lung cancer patients and 450 matched controls in this study sampled from the intervention arm.

Controls were individuals free of cancer at the time of a case's lung cancer diagnosis. Controls were individually matched to lung cancer patients on sex, date of birth +/- 1 year with a possible relaxation to 5 years, race, study year of blood draw, date of blood draw +/- 1 month (with a possible relaxation to 3/6 months), time of blood draw (6AM-9AM, 9AM-12PM, other), smoking categories (smoking status at enrollment, never, former, or current smoker; cumulative amount of smoking (0 to 29, 30 to 39, 40 to 49, and 50+ pack-years), with additional matching for time since quitting (< 15 years and ≥ 15 years) for former smokers, cigarettes per day smoked, and number of days in the study.

The New York University Women's Health Study (NYUWHS)

The New York University Women's Health Study (NYUWHS) is a prospective cohort study of women enrolled at a mammography screening center in New York City. From March 1985 through June 1991, 14,274 women between the ages of 34 and 65 were enrolled in the study. Because the original focus of the study was endogenous hormones and breast cancer, women who had taken hormone medications in the 6 months preceding baseline enrolment were not eligible for the study.

At the time of enrolment, data on demographics, anthropometric measures, medical history, reproductive and lifestyle variables were collected through self-administered questionnaires after written informed consent was obtained. Incident lung cancer cases were identified through active follow-up of the cohort conducted with questionnaires mailed approximately every two to four years and record linkages with state tumor registries in New York, New Jersey, and Florida, as well as record linkage with the National Death Index (NDI). Medical records were obtained to verify reported cancer outcomes. A total of 171 incident lung cancer cases and 171 individually matched controls were included for this analysis.

The American Cancer Society Cancer Prevention Study-II (CPS-II) Nutrition Cohort

The ACS Study-II is a prospective study of cancer incidence and mortality among 86,404 men and 97,786 women. The CPS-II Nutrition Cohort, which is described in detail elsewhere[5], was initiated in 1992 as a subgroup of CPS-II, a prospective study of cancer mortality involving approximately 1.2 million Americans begun in 1982. Participants in the CPS-II Nutrition Cohort were recruited from CPS-II members who resided in 21 states and were between the ages of 50 and 74 years. At enrollment in 1992/1993, participants completed a self-administered questionnaire that included demographic, medical, dietary, and lifestyle information. Follow-up questionnaires were sent to all living Nutrition Cohort members in 1997, and every two years after this to update exposure information and to ascertain newly diagnosed cancers. Between June 1998 and June 2001, blood samples were collected from a subset of CPS-II Nutrition Cohort participants (21,965 women and 17,411 men).

Incident lung cancer cases were identified through self-report on a follow-up questionnaire, linkage with state cancer registries, or death certificates. Self-reported cancers were verified through medical records. 200 incident lung cancer cases and 200 matched controls from the CPS-II Nutrition Cohort were provided for this analysis.

The Campaign Against Cancer and Stroke (CLUE I) and the Campaign Against Cancer and Heart Disease (CLUE II).

The CLUE studies include two large cohorts of volunteers from Washington County, Maryland that were enrolled in 1974 and 1989, respectively. CLUE I was conducted in Washington County, Maryland, in the fall of 1974. Brief health histories and blood pressures were taken and 15 ml of blood was drawn from 26,147 volunteers (23,951 were residents of Washington County) at the time of enrollment. Linkage of the records from this program to

those of a private census in the summer of 1975 indicated that almost a third of the adult population of the county had participated. CLUE II was an outgrowth of CLUE I conducted from May through October in 1989. As in CLUE I, a brief health history was obtained and 20 ml of blood was drawn. A blood sample was collected from 32,894 volunteers at the time of enrollment (25,076 were residents of Washington County). Participants were also given a food frequency questionnaire to complete at home and were asked to return it with a toenail clipping of the large toe for trace metal assays. Comparisons with published figures from the 1990 Census indicted that approximately 30 percent of adult residents had participated. For the current study, 15 case-sets were from CLUE1 and 185 were from CLUE2.

The Multiethnic Cohort (MEC)

The MEC includes over 215,000 men and women aged 45-75 years at recruitment from five different racial/ethnic groups (African Americans, Japanese Americans, Native Hawaiians, Latinos and European Americans) in Hawaii and California[6]. The cohort was assembled in 1993-1996 by mailing a self-administered, 26-page questionnaire to obtain extensive information on demographics, medical and reproductive histories, medication use, family history of various cancers, physical activity and diet. Identification of incident cancer cases is by regular linkage with the Hawaii, Los Angeles County and California SEER registries. From 1995 to 2001, blood collection was conducted from incident cases with breast, prostate, or colorectal cancers, as well as a random sample of cohort participants to serve as controls in genomic nested case-control studies (participation rate 72% and 63%, respectively). In addition, from 2001 to 2006, blood was also collected prospectively, without regard for cancer diagnosis, from willing cohort participants. Approximately 67,000 gave a blood sample (participation rate 43%). All incident lung cancer cases diagnosed before 2010 with a pre-diagnostic blood sample were considered for inclusion in this study. Each case was matched to a control based on study site, sex, age, race/ethnicity, smoking status, hours of fasting, and date and time of blood draw. For the current study, MEC provided 112 case-sets of white study participants and 62 case-sets of Asian study participants.

Women's Health Study (WHS)

The WHS was a randomized trial of low-dose aspirin, vitamin E, and beta-carotene in the primary prevention of cardiovascular disease and cancer beginning in 1992 among 39,876 female US health professionals aged ≥45 years[7]. Information on major clinical, lifestyle, and

dietary factors was collected via self-reports on baseline questionnaires. Women also provided baseline bloods. During more than two decades of follow-up, WHS participants reporting new cases of cancer on annual follow-up questionnaires were confirmed by medical record review by the WHS Endpoints Committee. Reports of cancer were confirmed on the basis of pathology or cytology reports or, rarely, strong clinical and radiologic or laboratory marker evidence when a pathology or cytology review was not conducted. Only confirmed cases of lung cancer were included in the present analyses, which were matched with eligible controls.

Physicians' Health Study (PHS)

The PHS I began in 1982 as a randomized trial of aspirin and beta-carotene for the primary prevention of heart disease and cancer among 22,071 male, Caucasian physicians initially aged 40 to 84 years[8], followed by the PHS II trial beginning in 1997 to evaluate beta-carotene, vitamin C, vitamin E, and a daily multivitamin on the prevention of cancer, CVD, and other endpoints. The PHS II included 14,641 men, with 7,641 participants from the PHS I plus 7,000 new physicians, for a total of 29,071 PHS participants[9]. A wide range of demographic, clinical, and lifestyle factors were assessed via baseline questionnaires, along with baseline bloods. PHS participants reported major clinical endpoints, including cancer, yearly in a mailed questionnaire and postcards every six months. Self-reported, incident lung cancer cases were confirmed through medical record review by the PHS Endpoints Committee in included in the present analyses.

The Nurses' Health Study (NHS)

The Nurses' Health Study (NHS)[10 11] was established in 1976, when 121,700 married female registered nurses aged 30 to 55 years residing in 11 States in the U.S. completed and returned a self-administered questionnaire. Questionnaires have been mailed to participants in both cohorts every 2 years since baseline to collect updated information on demographics, lifestyle factors, medical history, and disease outcomes. A semi quantitative food frequency questionnaire (FFQ) was administered to obtain information on usual dietary intake over the previous year. The reproducibility and validity of the FFQs have been established[12-16]. The FFQ was first administered in 1980 in the NHS, and were repeated almost every 4 years thereafter. For each food item, the questionnaire specified a common serving size and queried respondents on average intake during the previous year; responses in 9 categories ranged from almost never to 6 or more per day. Most nutritional variables

measured by these FFQs have been developed, tested, and refined by our group over the past 30 years (https://regepi.bwh.harvard.edu/health/).

The follow-up rate has been greater than 90%. The institutional review board at the Brigham and Women's Hospital approved the study. As approved by the committee, return of the questionnaires was considered to imply informed consent. Cases of lung cancer were self-reported by the participants or identified on their death certificates and were subsequently confirmed by medical records.

Health Professionnals Follow-up Study (HPFS)

The Health Professionals Follow-up Study (HPFS)[17] is an ongoing cohort study of 51,529 U.S. male professionals who were aged 40 to 75 years at baseline in 1986. Questionnaires have been mailed to participants in both cohorts every 2 years since baseline to collect updated information on demographics, lifestyle factors, mwedical history, and disease outcomes. The follow-up rate has been greater than 90%. The institutional review board at the Harvard T.H. Chan School Public Health approved this study. As approved by the committee, return of the questionnaires was considered to imply informed consent. A semiquantitative food frequency questionnaire (FFQ) was administered to obtain information on usual dietary intake over the previous year. The FFQ was first administered in 1986 in the HPFS and was repeated almost every 4 years thereafter. The reproducibility and validity of the FFQ have been estabished[14 18]. For each food item, the questionnaire specified a common serving size and queried respondents on average intake during the previous year; responses in 9 categories ranged from almost never to 6 or more per day. Cases of lung cancer were self-reported by the participants or identified on their death certificates and were subsequently confirmed by medical records.

We would like to thank the participants and staff of the Nurses' Health Study and the Health Professionals Follow-up Study, for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. This work was supported by the National Institutes of Health (NIH) grants, UM1CA167552, UM1CA186107, P01CA87969, and R01CA49449. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

2) European/Australian cohorts

The Melbourne Collaborative Cohort Study (MCCS)

The MCCS is a prospective cohort study of 41,514 participants (17,045 men and 24,469 women) aged 27-88 years at recruitment[19]; 99.3% of whom were aged 40-69 years. Recruitment occurred between 1990 and 1994. Southern European migrants to Australia (including 5,411 Italians and 4,525 Greeks) were over-sampled to extend the range of lifestyle exposures and to increase genetic variation.

Subjects were recruited via Electoral Rolls (registration to vote is compulsory for adults in Australia), advertisements, and community announcements in local media. Comprehensive lists of Italian and Greek surnames were used to target southern European migrants in phonebooks and electoral rolls. Passive follow-up of the cohort has been conducted by record linkage to Electoral Rolls, electronic phonebooks, the Victorian Cancer Registry and death records; as well as national cancer and death records to identify events outside of Victoria.

At recruitment participant's height and weight were measured, blood samples collected and questionnaires covering lifestyle (diet, smoking, physical activity and alcohol consumption), demographics and medical history completed.

Incident lung cancer cases under 80 years of age up to 31st December 2010 were identified through record linkage. A total of 361 lung cancer cases and 361 controls matched for date of birth, gender, smoking status, and ethnicity were included for this analysis.

The Malmö Diet and Cancer Study (MDCS)

The Malmö Diet and Cancer Study (MDCS) is a population-based prospective cohort study that between 1991 and 1996 recruited men and women aged 44 to 74 years of age living in Malmö, Sweden[20]. The main goal of the MDCS is to study the impact of diet on cancer incidence and mortality. It consists of a baseline examination including dietary assessment, a self-administered questionnaire, anthropometric measurements and collection of blood samples. A total of 201 incident lung cancer cases and 201 individually matched controls were available for this analysis.

The Northern Sweden Health and Disease Study Cohort (NSHDS)

The Northern Sweden Health and Disease Study (NSHDS) encompasses several prospective cohorts, the current study involving study participants from the Västerbotten Intervention Project (VIP), a sub-cohort within NSHDS[21]. VIP is an ongoing prospective cohort and

intervention study intended for health promotion of the general population of the Västerbotten County in northern Sweden. VIP was initiated in 1985 and all residents in the Västerbotten County were invited to participate by attending a health check-up at 40, 50 and 60 years of age. Participants were asked to complete a self-administered questionnaire including various demographic factors such as education, smoking habits, physical activity and diet. In addition, height and weight were measured and participants were asked to donate a fasting blood sample for future research. Incident lung cancer cases were identified through linkage with the regional cancer registry. A total of 245 incident lung cancer cases and 245 individually matched controls were included for this analysis.

The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC)

The ATBC Study was a randomized, double-blind, placebo-controlled, primary cancer prevention trial testing daily supplementation with α-tocopherol (50 mg/day) or β-carotene (20 mg/day), or both[22]. Between 1985 and 1988, the study enrolled and randomized 29,133 50-69 year old male cigarette smokers from southwestern Finland. Study supplementation continued for 5-8 years (median 6.1 years) until death or trial closure (April 30, 1993). At baseline, participants completed questionnaires regarding general risk factors, medical history, smoking habits, and dietary intake. Height, weight, heart rate, and blood pressure were measured by trained nurses and fasting serum samples were collected and stored at –70 °C. Lung cancer cases diagnosed during follow-up through 2009 were identified through linkage with the Finnish Cancer Registry, and 200 were individually matched to 200 controls for this analysis. For the current study, study participants were selected randomly from the different arms of the ATBC study.

The Nord-Trøndelag Health Study (HUNT)

The HUNT study is a longitudinal population based study having invited all persons aged 20-100 years living in the county of Nord-Trøndelag, Norway to three data collections, HUNT1 (1984-86), HUNT2 (1995-97) and HUNT 3 (2006-08) (http://www.ntnu.edu/hunt). Comprehensive data on life style, health status, symptoms, diseases and anthropometrics have been collected through questionnaires, interviews and clinical examinations, and in HUNT2 and HUNT3 biological material as blood and urine additionally were collected and stored. For this study, data have been linked to the Norwegian Cancer Registry and 238 incident lung cancer cases and 238 individually matched controls have been included.

3) Asian cohorts

The Shanghai Men's Health Study (SMHS) and the Shanghai Women's Health Study (SWHS)

The SMHS and SWHS are population-based cohort studies conducted in eight communities of urban Shanghai. Their designs and methods have been described elsewhere[23 24].

Briefly, the SWHS recruited 74,941 women during 1997-2000 (response rate: 93%) and the SMHS recruited 61,480 men during 2002-2006 (response rate: 74%). Similar methods and questionnaires were used in both studies. At baseline in-person interviews, information on sociodemographic, diet, lifestyle, occupation and medical history was obtained; height, body weight, and waist circumference were measured. Blood samples were collected from 75% of the study participants in both studies, processed within 6 hours, and stored at -70℃ until analysis.

The SMHS and SWHS have been followed up by annual record linkage with the population-based Shanghai Cancer Registry and Shanghai Vital Statistics Registry and in-person surveys every 2-3 years. Exposure information, including dietary intake, was updated in the in-person follow-up surveys. All possible matches from the linkages are checked manually and verified by home visits. Medical charts were obtained from the initial diagnostic hospitals to verify cancer diagnosis. Death certificate data from the Shanghai Vital Statistics Unit was used to identify the primary cause of death.

The studies were approved by the Institutional Review boards of the Shanghai Cancer Institute and Vanderbilt University. Informed consent was obtained from all participants.

The Singapore Chinese Health Study (SCHS)

The design of the SCHS study has been described[25 26]. Briefly, the cohort was drawn from permanent residents or citizens of Singapore who resided in government-built housing estates (86% of the Singapore population reside in such facilities). The eligible age range for cohort enrolment was 45-74 years. We restricted study subjects to the two major dialect groups of Chinese in Singapore: the Hokkiens and the Cantonese, who originated from Fujian and Guangdong provinces in Southern China, respectively. Between April 1993 and December 1998, 63,257 subjects (approximately 85% of eligible subjects) were enrolled into the cohort study. At recruitment, each study subject was interviewed in person by a trained interviewer using a structured questionnaire that emphasized current diet assessed via a validated, 165-item food frequency questionnaire. The questionnaire also requested information on demographics, lifetime use of tobacco, incense use, current physical activity,

usual sleep duration, reproductive history (women only), occupational exposure, medical history, and family history of cancer.

Beginning in April 1994, a random 3% sample of cohort participants were asked to provide blood or buccal cell (if request for blood sample was denied), and spot urine samples. Eligibility for this biospecimen subcohort was extended to all surviving cohort participants starting in January 2000. By April 2005, all surviving cohort subjects had been contacted for biospecimen donation. Approximately 60% of eligible cohort participants donated biospecimens.

The cohort has been passively followed for death and cancer occurrence through regular record linkage with the population-based Singapore Cancer Registry and the Singapore Registry of Births and Deaths. Migration out of Singapore, especially among housing estate residents, is negligible. As of latest update, only 55 individuals from this cohort were known to be lost to follow-up due to migration and other reason.

A nested case-control study of incident lung cancer cases within the Singapore Chinese Health Study was used to examine the association between serum levels of vitamin B_6 and other compounds in the one-carbon metabolism pathway and risk of lung cancer. Briefly, 422 lung cancer cases were identified among cohort participants with available prediagnostic plasma samples as of 12/31/2011. For each case, one control subject was randomly selected from all eligible cohort members who were alive and free of cancer on the date of cancer diagnosis of the index case. The control subject was individually matched to the index case by gender, dialect group (Hokkien, Cantonese), age at enrolment (±3 years), date of baseline interview (±2 year), date of biospecimen collection (±6 months), and smoking status (current, former, and never smokers). For current smokers, cases and controls were further matched by number of cigarettes per day (<15, ≥15 cigarettes/day). For former smokers, cases and controls were further matched by years since quitting smoking (<10, ≥10 years). One plasma aliquot per subject was retrieved from the biorepository and all plasma samples were sent to the laboratory (B-vital) for measurements.

The Shanghai Cohort Study (SCS)

The SCS study is a residential cohort of 18,244 men in Shanghai, China, assembled during 1986-89 when subjects were between the ages of 45 and 64 years. Approximately 80% of

eligible men participated in the study. At the time of recruitment, each cohort subject was interviewed in-person by a trained nurse interviewer using a structured questionnaire that included background information, history of tobacco and alcohol use, current diet, and medical history[27 28].

At the completion of the interview, the nurse collected a 10 ml blood and a single void urine specimen from the study participant. Blood and urine samples were kept in insulated boxes with ice (0-2^{o}C). The serum was separated from blood specimen within 3-4 hours after collection. Two sets of serum (2 ml and 1 ml, respectively) and two sets of urine samples (10 ml each) per subject have been stored at -80^{o}C.

The cohort has been followed for the occurrence of cancer and death through routine ascertainment of new cases from the population-based Shanghai Cancer Registry and Shanghai Vital Statistics Units. To maximize the cancer findings and minimize the loss of follow-up, we have recontacted each surviving cohort member annually. Retired nurses visit the last known address of each living cohort member and record details of the interim health history of the cohort member. As of December 31, 2014, cumulatively 612 (3.4%) original subjects were lost to follow-up (i.e., persons we have no record of death and we have been unable to locate through our annual follow-up recontacts), and 574 (3.1%) refused to our continued follow-up interview (their cancer and vital status has been continually updated through record linkage analyses) after 26 years of follow-up since the beginning of the study.

A nested case-control study of incident lung cancer cases within the Shanghai Cohort Study was used to examine the association between serum levels of vitamin B_6 and other compounds in the one-carbon metabolism pathway and risk of lung cancer. Briefly, 516 lung cancer cases were identified among cohort participants with available serum samples as of 12/31/2006. For each case, we randomly selected one control subject from all cohort members who were free of cancer and alive at the time of cancer diagnosis of the index case. Controls were matched to the index case by age at enrolment (±2 years), date of biospecimen collection (±1 month) and neighbourhood of residence at recruitment, and smoking status (current, former and never smokers) as established previously for other studies. For former smokers, cases and controls were further matched by years since

quitting smoking (<10 vs ≥10 years). One serum vial per subject was retrieved from biorepository and all serum samples were sent to the laboratory (B-vital) for measurements.

Biochemical analyses

Serum or plasma samples from all study participants were sent on dry ice to the Bevital A/S laboratory (http://www.bevital.no) in Bergen, Norway, were they were kept below −80°C until analysis. Concentrations of vitamin B6, methionine and cotinine were determined by mass spectrometry based methods (LC-MS/MS, GC-MS/MS),[29, 30], and microbiological methods were used to determine concentrations of folate (Lactobacillus casei).[31] Samples were analyzed in batches of 86 samples, and quality control procedures included 6 calibration samples, 2 control samples, and 1 blank sample in each batch. The coefficients of variation (CVs) within and between batches were, respectively, 3% and 7% for vitamin B6, 4% and 5% for folate, and 1% and 3% for methionine. All lung cancer cases and their individually matched controls were analyzed together within the same batches in random order. The laboratory staff was blinded to the case-control status of the blood samples.

Statistical analyses

We fitted hierarchical linear models to the biomarker concentrations to evaluate between-cohort variability in mean concentrations, and the extent to which this variability could be accounted for by individual level covariates such as age, sex, body mass index, and smoking status. To account for the nested nature of the data (with participants nested within cohorts) and to allow model parameters to vary across cohorts, we used a hierarchical linear regression model to regress standardized, log transformed biomarker concentrations on these covariates. In particular, let $\mathbf{y}$ denote the outcome variable (i.e., log transformed concentrations), and X denote a matrix of K covariates. We model the outcome of participant n from cohort j as

$$y_{nj} \sim \text{Normal}(\alpha_j + x_{nj}\boldsymbol{\beta}_j, \sigma^2).$$

Each of the K components of the cohort-specific vector of regression coefficients β_{jk} are modeled with independent normal distributions, as are the cohort-specific intercepts α_j:

$$\beta_{jk} \sim \text{Normal}(\mu_k, \tau_k^2);\ \alpha_j \sim \text{Normal}(\phi, \delta^2).$$

The overall intercept ϕ and mean regression coefficients μ_k were given diffuse $\text{Normal}(0, 50^2)$ prior distributions, and the hierarchical standard deviation parameters τ_k and δ were given weakly informative $\text{HalfCauchy}(0,2)$ hyperprior distributions. The residual

standard deviation σ was also given a $\mathrm{HalfCauchy}(0,2)$ prior distribution. We conducted sensitivity analyses in which we specified different scales for the Half-Cauchy hyperpriors, and found that the posterior distributions of α_j and β_{jk} were insensitive to the specified scale of the hyperpriors.

Risk analysis involved calculating quartiles of circulating concentrations for each biomarker based on the distribution among controls, both overall and by region (United States [US], Europe and Australia combined [EU/AU], and Asia). Relative risks of lung cancer were estimated by calculating odds ratios (OR) and 95% confidence intervals (CI) using conditional logistic regression (conditioning on matched case-control sets) with the first quartile as the referent. Additional covariates were included to account for confounding by risk factors, including indicators of educational attainment (in six categories) and tobacco exposure (in addition to matching to smoking status in 5 categories: cotinine concentrations [defined using quartiles of the distribution among participants reporting to be current smokers]). Including additional covariates of body mass index (BMI) and alcohol intake did not appreciably alter the results and were not included in the final models. As a sensitivity analysis, we fitted models that were additionally adjusted for smoking duration or pack-years of smoking among ever smokers, 90% of which had such information available. All risk analyses were conducted overall and by stratifying for never, former and current smokers, as well as by region (US, EU/AU, Asia).

As an indication of the overall statistical strength of association between each biomarker and risk, we calculated a p-value for trend by including the base-2 logarithm (log2) of the biomarker concentration as a continuous variable in a separate conditional logistic regression model. The same approach was used in stratified risk analyses according to other pre-defined demographic characteristics and risk factors. OR estimates per log2 unit ($\log_2$ OR) may be interpreted as the relative risk associated with a doubling in the concentration of a circulating biomarker. We also performed risk analyses for $\log_2$OR on each biomarker separately for each individual cohort, as well for US cohorts by the period of blood sample collection (before and after 1996 when folate fortification of food items took effect) to evaluate potential impact of folate fortification. X^2 tests were used to assess heterogeneity in $\log_2$ OR estimates in stratified analyses.

Finally, in order to evaluate the specific hypothesis of whether deficiency in vitamin B6 or folate is particularly important in lung cancer etiology, we conducted analysis by

categorizing subjects into deficient, or normal category based on clinically defined cut-off values. For vitamin B6, <20 nmol/L was considered deficient whereas ≥20nmol/L was considered normal [32]. For folate deficiency, subjects were classified in three categories; <7 nmol/L as deficient, 7-12.9 nmol/L as moderately deficient, and ≥ 13 nmol/L as normal [33 34]. All statistical analyses were conducted using SAS 9.2 (Cary, North Carolina) [35], R version 3.1.3 [36] or Stan version 2.9.0 [37].

References

1. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials* 1998;19(1):61-109.
2. Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, et al. The Women's Health Initiative recruitment methods and results. *Ann Epidemiol* 2003;13(9 Suppl):S18-77.
3. Signorello LB, Hargreaves MK, Steinwandel MD, Zheng W, Cai Q, Schlundt DG, et al. Southern community cohort study: establishing a cohort to investigate health disparities. *J Natl Med Assoc* 2005;97(7):972-9.
4. Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 2000;21(6 Suppl):273S-309S.
5. Calle EE, Rodriguez C, Jacobs EJ, Almon ML, Chao A, McCullough ML, et al. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer* 2002;94(2):500-11.
6. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am J Epidemiol* 2000;151(4):346-57.
7. Buring JE HC. The women's health study: Rationale and background. *J Myocardial Ischemia* 1992;4:30-40.
8. Final Report on the Aspirin Component of the Ongoing Physician's Health Study. *NEJM* 1989 July;321:129-35.
9. Christen WG, Gaziano JM, Hennekens CH. Design of Physicians' Health Study II--a randomized trial of beta-carotene, vitamins E and C, and multivitamins, in prevention of cancer, cardiovascular disease, and eye disease, and review of results of completed trials. *Ann Epidemiol* 2000;10(2):125-34.
10. Belanger CF, Hennekens CH, Rosner B, Speizer FE. The nurses' health study. *Am. J. Nurs.* 1978;78(6):1039-40.
11. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer* 2005;5(5):388-96.
12. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 1985;122(1):51-65.
13. Wu K, Willett WC, Fuchs CS, Colditz GA, Giovannucci EL. Calcium intake and risk of colon cancer in women and men. *J Natl Cancer Inst* 2002;94(6):437-46.
14. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol* 1992;135(10):1114-26; discussion 27-36.
15. Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B, et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. *Int. J. Epidemiol.* 1989;18(4):858-67.

16. Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J. Am. Diet. Assoc.* 1993;93:790-96.
17. Giovannucci E, Ascherio A, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Physical activity, obesity, and risk for colon cancer and adenoma in men. *Ann Intern Med* 1995;122(5):327-34.
18. Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J Am Diet Assoc* 1993;93(7):790-6.
19. Giles GG, English DR. The Melbourne Collaborative Cohort Study. *IARC Sci Publ* 2002;156:69-70.
20. Berglund G, Elmstahl S, Janzon L, Larsson SA. The Malmo Diet and Cancer Study. Design and feasibility. *J Intern Med* 1993;233(1):45-51.
21. Hallmans G, Agren A, Johansson G, Johansson A, Stegmayr B, Jansson JH, et al. Cardiovascular disease and diabetes in the Northern Sweden Health and Disease Study Cohort - evaluation of risk factors and their interactions. *Scand J Public Health Suppl* 2003;61:18-24.
22. The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. The ATBC Cancer Prevention Study Group. *Ann Epidemiol* 1994;4(1):1-10.
23. Shu XO, Li H, Yang G, Gao J, Cai H, Takata Y, et al. Cohort Profile: The Shanghai Men's Health Study. *Int J Epidemiol* 2015;44(3):810-8.
24. Zheng W, Chow WH, Yang G, Jin F, Rothman N, Blair A, et al. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. *Am J Epidemiol* 2005;162(11):1123-31.
25. Hankin JH, Stram DO, Arakawa K, Park S, Low SH, Lee HP, et al. Singapore Chinese Health Study: development, validation, and calibration of the quantitative food frequency questionnaire. *Nutr Cancer* 2001;39(2):187-95.
26. Yuan JM, Stram DO, Arakawa K, Lee HP, Yu MC. Dietary cryptoxanthin and reduced risk of lung cancer: the Singapore Chinese Health Study. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2003;12(9):890-8.
27. Ross RK, Yuan JM, Yu MC, Wogan GN, Qian GS, Tu JT, et al. Urinary aflatoxin biomarkers and risk of hepatocellular carcinoma. *Lancet* 1992;339(8799):943-6.
28. Yuan JM, Ross RK, Wang XL, Gao YT, Henderson BE, Yu MC. Morbidity and mortality in relation to cigarette smoking in Shanghai, China. A prospective male cohort study. *Jama* 1996;275(21):1646-50.
29. Midttun O, Hustad S, Ueland PM. Quantitative profiling of biomarkers related to B-vitamin status, tryptophan metabolism and inflammation in human plasma by liquid chromatography/tandem mass spectrometry. *Rapid communications in mass spectrometry : RCM* 2009;23(9):1371-9.
30. Ueland PM, Midttun O, Windelberg A, Svardal A, Skalevik R, Hustad S. Quantitative profiling of folate and one-carbon metabolism in large-scale epidemiological studies by mass spectrometry. *Clinical chemistry and laboratory medicine : CCLM / FESCC* 2007;45(12):1737-45.
31. Molloy AM, Scott JM. Microbiological assay for serum, plasma, and red cell folate using cryopreserved, microtiter plate method. *Methods in enzymology* 1997;281:43-53.
32. Ueland PM, Ulvik A, Rios-Avila L, Midttun O, Gregory JF. Direct and Functional Biomarkers of Vitamin B6 Status. *Annu Rev Nutr* 2015;35:33-70.
33. de Benoist B. Conclusions of a WHO Technical Consultation on folate and vitamin B12 deficiencies. *Food and nutrition bulletin* 2008;29(2 Suppl):S238-44.
34. Bailey LB, Stover PJ, McNulty H, Fenech MF, Gregory JF, 3rd, Mills JL, et al. Biomarkers of Nutrition for Development-Folate Review. *The Journal of nutrition* 2015;145(7):1636S-80S.
35. SAS® 9.4 Help and documentation [program]. 9.4 version. Cary, NC, USA: SAS Institute Inc, 2014.

36. RStan: the R interface to Stan [program]. Version 2.5 version, 2014.
37. Stan Modeling Language Users Guide and Reference Manual [program]. Version 2.5.0. version, 2014.