Eur J Epidemiol (2013) 28:677–688
DOI 10.1007/s10654-013-9793-z

CANCER

A structural equation modelling approach to explore the role of B vitamins and immune markers in lung cancer risk

Valéria Troncoso Baltar · Wei W. Xun · Mattias Johansson · Pietro Ferrari · Shu-Chun Chuang · Caroline Relton · Per Magne Ueland · Øivind Midttun · Nadia Slimani · Mazda Jenab · Françoise Clavel-Chapelon · Marie-Christine Boutron-Ruault · Guy Fagherazzi · Rudolf Kaaks · Sabine Rohrmann · Heiner Boeing · Cornelia Weikert · Bas Bueno-de-Mesquita · Hendriek Boshuizen · Carla H. van Gils · N. Charlotte Onland-Moret · Antonio Agudo · Aurelio Barricarte · Carmen Navarro · Laudina Rodríguez · José Maria Huerta Castaño · Nerea Larrañaga · Kay-Tee Khaw · Nick Wareham · Naomi E. Allen · Francesca Crowe · Valentina Gallo · Teresa Norat · Vittorio Krogh · Giovanna Masala · Salvatore Panico · Carlotta Sacerdote · Rosario Tumino · Antonia Trichopoulou · Pagona Lagiou · Dimitrios Trichopoulos · Torgny Rasmuson · Göran Hallmans · Nina Roswall · Anne Tjønneland · Elio Riboli · Paul Brennan · Paolo Vineis

Received: 27 August 2012 / Accepted: 1 March 2013 / Published online: 27 March 2013

Abstract The one-carbon metabolism (OCM) is considered key in maintaining DNA integrity and regulating gene expression, and may be involved in the process of carcinogenesis. Several B-vitamins and amino acids have been implicated in lung cancer risk, via the OCM directly as well as immune system activation. However it is unclear whether these factors act independently or through complex mechanisms. The current study applies structural equations modelling (SEM) to further disentangle the mechanisms involved in lung carcinogenesis. SEM allows

V. T. Baltar (✉)
Department of Epidemiology, Faculty of Public Health, University of São Paulo, Av. Dr Arnaldo, 715, São Paulo, São Paulo CEP 01246-904, Brazil
e-mail: vbaltar@usp.br

V. T. Baltar · W. W. Xun · S.-C. Chuang · V. Gallo · T. Norat · E. Riboli
Department of Epidemiology and Biostatistics, Faculty of Medicine, School of Public Health, Imperial College London, London, UK
e-mail: wei.xun06@imperial.ac.uk

S.-C. Chuang
e-mail: s-c.chuang@imperial.ac.uk

V. Gallo
e-mail: v.gallo@imperial.ac.uk

T. Norat
e-mail: t.norat@imperial.ac.uk

E. Riboli
e-mail: e.riboli@imperial.ac.uk

M. Johansson · P. Ferrari · N. Slimani · M. Jenab · P. Brennan
International Agency for Research on Cancer, Lyon, France
e-mail: slimani@iarc.fr

M. Ferrari
e-mail: ferrarip@iarc.fr

N. Slimani
e-mail: slimani@iarc.fr

M. Jenab
e-mail: jenab@iarc.fr

P. Brennan
e-mail: brennan@iarc.fr

C. Relton
Institute for Ageing and Health, Newcastle University, Campus for Ageing and Vitality, Newcastle, UK
e-mail: c.l.relton@newcastle.ac.uk

P. M. Ueland
Section for Pharmacology, Institute of Medicine, University of Bergen, Bergen, Norway
e-mail: per.ueland@ikb.uib.no

P. M. Ueland
Haukeland University Hospital, Bergen, Norway

simultaneous estimation of linear relations where a variable can be the outcome in one equation and the predictor in another, as well as allowing estimation using latent variables (factors estimated by correlation matrix). A large number of biomarkers have been analysed from 891 lung cancer cases and 1,747 controls nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Four putative mechanisms in the OCM and immunity were investigated in relation to lung cancer risk: methionine-homocysteine metabolism, folate cycle, transsulfuration, and mechanisms involved in inflammation and immune activation, all adjusted for tobacco exposure. The hypothesized SEM model confirmed a direct and protective effect for factors representing methionine-homocysteine metabolism ($p = 0.020$) and immune activation ($p = 0.021$), and an indirect protective effect of folate cycle ($p = 0.019$), after adjustment for tobacco smoking. In conclusion, our results show that in the investigation of the involvement of the OCM, the folate cycle and immune system in lung carcinogenesis, it is important to consider complex pathways (by applying SEM) rather than the effects of single vitamins or nutrients (e.g. using traditional multiple regression). In our study SEM were able to suggest a greater role of the methionine-homocysteine metabolism and immune activation over other potential mechanisms.

Keywords B vitamins · Folate · Methionine · Lung cancer · Immune markers · Structural equation model

Introduction

Lung cancer continues to be a major public health problem world-wide. The World Health Organization estimated that in 2008 1.37 million deaths were due to cancer of the lung, out of 7.6 million deaths from all cancers [1]. Reducing tobacco consumption remains the most appropriate harm reduction strategy [1, 2]. In addition to exposures to carcinogenic mixtures such as smoking and air pollution, low-micronutrient intake has been linked to degenerative diseases including lung cancer [3, 4].

Amongst the micronutrients that have been investigated, certain B vitamins and other one-carbon metabolism (OCM) components have been associated with carcinogenesis [3, 4]. The OCM describes a network of several reactions including the folate (vitamin B9) cycle, in which folate acts as a crucial carrier of methyl groups and involves several additional water-soluble B vitamins such as B6 and B12, as well as homocysteine, methionine and serine (see Figure 1 of Johansson et al. [5]), most of which are derived from diet. One branch of the metabolism promotes the synthesis of methionine and S-adenosylmethionine (SAM)

Ø. Midttun
Bevital AS, Bergen, Norway
e-mail: bjorn.midttun@farm.uib.no

F. Clavel-Chapelon · M.-C. Boutron-Ruault · G. Fagherazzi
Institut National de la Santé et de la Recherche Médicale (INSERM) U780-IFR169, Villejuif, France
e-mail: clavel@igr.fr

M.-C. Boutron-Ruault
e-mail: boutron@igr.fr

G. Fagherazzi
e-mail: guy.fagherazzi@igr.fr

F. Clavel-Chapelon · M.-C. Boutron-Ruault · G. Fagherazzi
Institut Gustave Roussy, Villejuif, France

F. Clavel-Chapelon · M.-C. Boutron-Ruault · G. Fagherazzi
Paris South University, Villejuif, France

R. Kaaks
Division of Cancer Epidemiology, German Cancer Research Centre, Heidelberg, Germany
e-mail: r.kaaks@dkfz-heidelberg.de

S. Rohrmann
Division of Clinical Epidemiology, German Cancer Research Centre, Heidelberg, Germany
e-mail: sabine.rohrmann@ifspm.uzh.ch

H. Boeing · C. Weikert
Department of Epidemiology, German Institute of Human Nutrition, Nuthetal, Germany
e-mail: boeing@mail.dife.de

C. Weikert
e-mail: weikert@dife.de

B. Bueno-de-Mesquita
Department for Determinants of Chronic Diseases, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
e-mail: Bas.Bueno.De.Mesquita@rivm.nl

H. Boshuizen
Department of Epidemiology, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
e-mail: Hendriek.Boshuizen@rivm.nl

C. H. van Gils · N. C. Onland-Moret
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
e-mail: C.vanGils@umcutrecht.nl

N. C. Onland-Moret
e-mail: N.C.Onland@umcutrecht.nl

which is essential for DNA methylation. This is modulated by the intake of dietary folate, which acts as a carrier of single-carbon units, essential for both the biosynthesis of nucleotides such as purines and pyrimidines, as well as in the process of remethylation of homocysteine to methionine [6–10]. Folate is metabolised to 5, 10-methylenetetrahydrofolate, which is subsequently converted by the enzyme methylenetetrahydrofolate reductase (MTHFR) to 5-methyltetrahydrofolate, by adding a methyl group to homocysteine to produce methionine and SAM [11]. Vitamin B12 is essential in this conversion [12]. Folate or vitamin B12 deficiency results in decreased SAM which, in turn, leads to elevated homocysteine levels [8], due to decreased methionine synthesis. Finally, SAM with the enzyme DNA methyltransferase leads to DNA methylation. During DNA methylation some methionine is re-used for homocysteine remethylation [13].

During transsulfuration, vitamin B6 is a cofactor for the conversion from homocysteine to cystathionine, and subsequently to cysteine [14]. Therefore, homocysteine can be irreversibly catabolized during transsulfuration. An inverse relation of B6 and homocysteine has been reported previously [15–17].

Vitamin B6 acts as a cofactor in more than 100 mechanisms in the body [18], including the transsulfuration pathway [19], and several steps of the kynurenine pathway of tryptophan metabolism that are related to inflammation and immune activation [20]. Tryptophan is obtained from diet and is transformed to kynurenine. The depletion of tryptophan and increased kynurenine-tryptophan ratio (KTR) have been reported in viral, bacterial and parasitic infections [21–24], and also in induced inflammation [25]. Furthermore, increased neopterin has been reported in patients with activated cellular immune response such as viral infections, autoimmune disorders, cardiovascular disease and cancer [26]. Neopterin seems part of the pro-inflammatory and cytocidal armature of the activated immune system [27]. Sucher et al. [28] discuss neopterin as an indicator of systemic immune activation in patients with malignant disease.

Two primary mechanisms by which OCM has been proposed to directly influence the process of carcinogenesis are: (1) DNA methylation via methionine-homocysteine metabolism, and (2) DNA synthesis and repair via the folate cycle [6, 8]. The transsulfuration mechanism may indirectly influence carcinogenesis by its role in homocysteine catabolism.

The addition of a methyl group at the carbon-5' position of the cytosine ring is termed DNA methylation, a frequent epigenetic event. It is an important regulatory mechanism in gene transcription and can initiate processes which have been implicated in a variety of tumours [12]. Several genes were reported as hyper-methylated in human lung cancer [12, 29, 30]. Vineis et al. [31] found a positive relation between serum folate and methylation levels of genes RASSF1A and MTHFR in never-smokers. Folate deficiency can also cause thymidylate stress during DNA synthesis,

A. Agudo
Unit of Nutrition, Environment and Cancer, Catalan Institute of Oncology, L'Hospitalet de Llobregat, Barcelona, Spain
e-mail: a.agudo@iconcologia.net

A. Barricarte
Epidemiology, Prevention and Promotion Health Service, Public Health Institute of Navarra, Navarra, Spain
e-mail: abarricg@cfnavarra.es

A. Barricarte · C. Navarro · J. M. H. Castaño · N. Larrañaga
CIBER of Epidemiology and Public Health, Barcelona, Spain
e-mail: carmen.navarro@carm.es

J. M. H. Castaño
e-mail: jmhuerta.carm@gmail.com

N. Larrañaga
e-mail: epidem3-san@ej-gv.es

C. Navarro
Department of Epidemiology, Regional Council of Health and Consumer Affairs, Murcia, Spain

L. Rodríguez
Public Health and Participation Directorate, Health and Health Care Services Council, Asturias, Spain
e-mail: laudina.rodriguezsuarez@asturias.org

J. M. H. Castaño
Department of Epidemiology, Murcia Regional Health Authority, Murcia, Spain

N. Larrañaga
Epidemiology Unit, Public Health Department of Gipuzkoa, Basque Government, San Sebastián, Spain

K.-T. Khaw
Clinical Gerontology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
e-mail: kk101@medschl.cam.ac.uk

N. Wareham
MRC Epidemiology Unit, Strangeways Research Laboratory, Cambridge, UK
e-mail: nick.wareham@mrc-epid.cam.ac.uk

N. E. Allen · F. Crowe
Cancer Epidemiology Unit, University of Oxford, Oxford, UK
e-mail: Naomi.allen@ceu.ox.ac.uk

F. Crowe
e-mail: francesca.crowe@ceu.ox.ac.uk

thus resulting in DNA strand breaks and genomic instability [6, 12].

In a recent case–control study of lung cancer nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) study, protective effects were observed for increasing serum levels of vitamin B6 and methionine, after adjustments for tobacco smoking and other potential confounders [8]. In adjusted logistic regression analysis, these effects were considered independently, but it is possible that they interact in a more complex way in their relation to cancer risk, which is not straightforwardly captured by standard regression analysis.

In order to further investigate the complex network through which OCM components and immune activation might influence lung cancer risk, we have conducted a study using the EPIC lung cancer data set [8]. We applied structural equation models (SEM) to measure four latent variables directly or indirectly related to lung cancer risk. Each latent variable represented one of the putative mechanisms of carcinogenesis: the methionine—homocysteine metabolism (DNA methylation), the folate cycle (DNA synthesis), transsulfuration, and immune activation. The hypothesized pathways were constructed assuming that groups of specific biomarkers are inter-related and organised into different latent factors (i.e. they cannot be measured directly), which represent the underlying mechanisms mentioned above. As tobacco exposure (measured as a combination of smoking status and cotinine levels at recruitment) is the chief risk factor for lung cancer, we adjusted the direct effect to lung cancer for this exposure and we also included it as indirect effect through the factors.

Materials and methods

Study subjects

A nested case–control study on lung cancer has been set up in the EPIC cohort. The EPIC study is a multi-centre study in 10 European Countries and 23 centres [32], involving more than 520,000 participants and with 10 years of follow-up. The lung cancer cases were defined on the basis of the International Classification of Diseases for Oncology, Second Edition, and included all invasive cancers coded as C34. Cases with no available blood sample or date of blood collection or history of another cancer at the time of blood donation were excluded. For each case, two controls were randomly matched considering: country, gender, date of blood collection and date of birth [8]. Another 39 subjects were excluded because of missing information from laboratory analyses or because their matched counterpart(s) had missing information. The final data consisted of 2,638 subjects (891 cases and 1,747 controls: 856 cases with 1:2 matching and 35 with 1:1). The study was approved by local ethics committees.

Biochemical analyses

Measurements of serum concentrations at recruitment (plasma in Swedish samples) of serine, glycine, vitamin B6 (PLP), folate, vitamin B12 (Cobalamin), vitamin B2 (riboflavin), total homocysteine, methionine, cysteine, cystathionine, neopterin, kynurenine/tryptophan ratio and cotinine were performed at

V. Krogh
Nutritional Epidemiology Unit, National Cancer Institute, Milan, Italy
e-mail: vittorio.krogh@istitutotumori.mi.it

G. Masala
Molecular and Nutritional Epidemiology Unit, ISPO—Cancer Research and Prevention Institute, Florence, Italy
e-mail: g.masala@ispo.toscana.it

S. Panico
Department of Clinical and Experimental Medicine, Federico II University of Naples, Naples, Italy
e-mail: spanico@unina.it

C. Sacerdote · P. Vineis
Human Genetics Foundation (HuGeF), Turin, Italy
e-mail: carlotta.sacerdote@cpo.it

P. Vineis
e-mail: p.vineis@imperial.ac.uk

R. Tumino
Cancer Registry and Histopathology Unit, "Civile M.P.Arezzo" Hospital, ASP 7, Ragusa, Italy
e-mail: rtumino@tin.it

A. Trichopoulou · P. Lagiou
Department of Hygiene, Epidemiology and Medical Statistics, WHO Collaborating Center for Food and Nutrition Policies, University of Athens Medical School, Athens, Greece
e-mail: atrichopoulou@hhf-greece.gr

P. Lagiou
e-mail: pdlagiou@med.uoa.gr

A. Trichopoulou
Hellenic Health Foundation, Athens, Greece

P. Lagiou · D. Trichopoulos
Department of Epidemiology, Harvard School of Public Health, Harvard University, Boston, MA, USA
e-mail: dtrichop@hsph.harvard.edu

D. Trichopoulos
Bureau of Epidemiologic Research, Academy of Athens, Athens, Greece

T. Rasmuson
Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden
e-mail: torgny.rasmuson@onkologi.umu.se

Bevital A/S (www.bevital.no), blinded as to the case–control status [8].

Model assignment and statistical methods

SEM with latent variables were applied. A latent variable (referred as *factor* from here on) is defined as a random variable which cannot be measured directly. The basic assumption is that the co-variation observed among the measured variables (i.e. the B vitamins, methionine etc.) is due to their relationship with the factors [33]. Therefore, a factor is estimated on the basis of observed variables [34].

On the basis of the background literature, we constructed four factors which were linked to lung cancer status; the observed biomarkers which make up each factor are as follows:

Factor 1 *Methionine-homocysteine metabolism* methionine, homocysteine, and Vitamin B12;
Factor 2 *Folate cycle* Vitamin B6, folate, vitamin B2 and glycine;
Factor 3 *Inflammation and Immune activation* Vitamin B6, neopterin and KTR;
Factor 4 *Transsulfuration* homocysteine, Vitamin B6, Cystathionine and Cysteine

Due to expected strong influence of folate cycle on methionine-homocysteine metabolism [13], we hypothesized an indirect effect of factor 2 on lung cancer risk through factor 1. We also hypothesized Factor 4 (transsulfuration) as an indirect effect on lung cancer risk through factor 1 (methionine-homocysteine metabolism), since the disruption of transsulfuration can increase homocysteine levels. Since smoking status and level of tobacco exposure are important causes of lung cancer, we adjusted all regressions for tobacco exposure, using a composite variable combining smoking status and cotinine levels in five categories: (a) never smokers and cotinine levels lower than 5 nmol/L; (b) former smokers and cotinine levels lower than 5 nmol/L; (c) never/former smokers with cotinine levels between 5 and 85 nmol/L; (d) cotinine level between 85 and 1700 nmol/L; (e) cotinine levels higher than 1,700 nmol/L [35]. This variable combines smoking status with passive and active exposure (short-term).

We have started the modeling with the hypothesized model represented in Fig. 1 (each square represents an observed variable, whereas each circle represents a factor; each arrow represents one causal direction). This model included four factors. Factor 2 (folate cycle) was considered as a direct and an indirect risk factor of lung cancer whereas factor 4 (transsulfuration) was considered only as an indirect risk factor. This first model, however, showed very low goodness-of-fit (see Table 2). Since among the four factors, transsulfuration lacks an established a priori role in lung cancer risk, it was excluded. A second model was hypothesized with three factors (folate cycle, methionine-homocysteine and Inflammation and Immune activation) in which factor 2 was considered as a direct and an indirect risk factor of lung cancer (through factor 1). From this point several models were simulated considering inclusion or exclusion of each component for each factor, maintaining the hypothesized general pathway of the second model with and without direct effect of factor 2. We presented as final model the one with the best goodness-of-fit (Table 2; Fig. 2).

All analyses were conducted in Mplus version 6, which is an up-to-date software specializing in SEM, in which it is possible to model continuous and binary outcomes for matched samples [36]. For the final binary outcome we considered the cases and controls in a *probit* SEM model. The *probit* model is used to model dichotomous outcome variables and the inverse standard normal distribution of the probability is modelled as a linear combination of the predictors. To consider the matching design we used the weighted least square method with mean and variance adjusted (WLSMV) estimation for SEM model estimation [36, 37]. This method also has the chief advantage of making minimal assumptions about the distribution of the observed variables. All observed variables were previously standardized (minus sample mean and divided by standard deviation) and considered with measurement errors represented by ε's (Fig. 2).

The following standard SEM adequacy fit indices were used to assess model fit: the comparative fit index (CFI; >0.90 considered as adequate), the Tucker-Lewis index (TLI; >0.90 considered as adequate), the root mean square error of approximation (RMSEA; <0.05 = good, 0.05–0.08 = adequate, 0.08–0.10 = marginal, >0.10 = poor) and Chi square [38]. The Chi square is an absolute fit index, TLI is a relative fit index and RMSEA and CFI complete the analysis as noncentrality-based indices.

G. Hallmans
Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden
e-mail: goran.hallmans@nutrires.umu.se

N. Roswall · A. Tjønneland
Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark
e-mail: roswall@cancer.dk

A. Tjønneland
e-mail: annet@cancer.dk

P. Vineis
Department of Epidemiology and Biostatistics, MRC-HPA Centre for Environment and Health, Imperial College London, London, UK

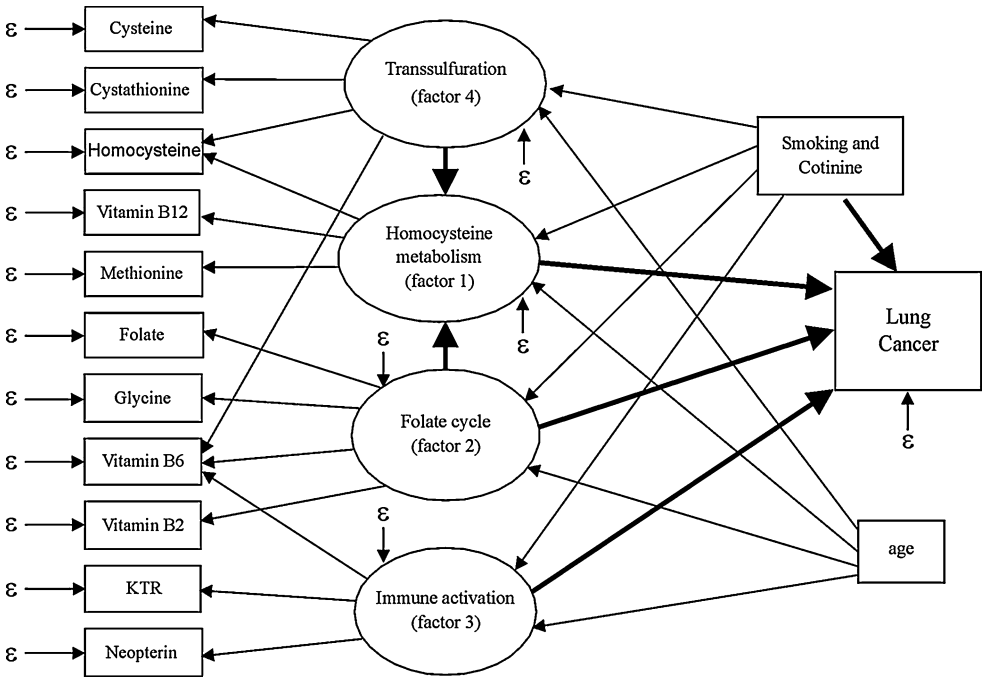

Fig. 1 Initial hypothesized model representing Methionine-homocysteine metabolism, folate cycle, transsulfuration and immune activation pathways and their effect on lung cancer risk

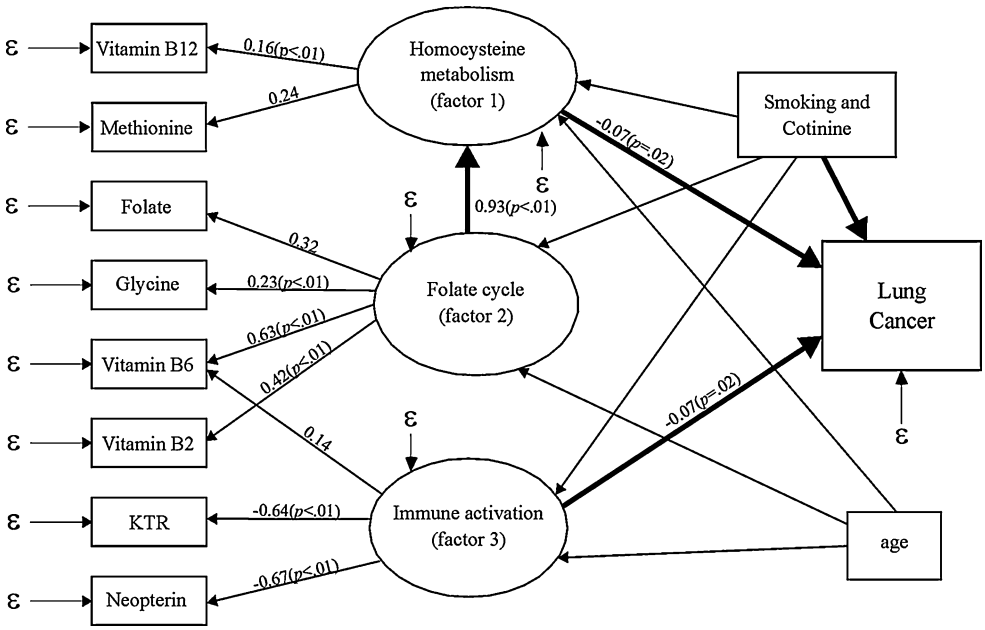

Fig. 2 Final hypothesized model representing Methionine-homocysteine metabolism, folate cycle and immune activation pathways and their effect on lung cancer risk

Results

In the current nested case–control study, the mean age was 58 years for both cases and controls. Amongst controls, 31.9 % were never smokers with cotinine levels lower than 5 nmol/L, 36.4 % were former smokers with cotinine levels lower than 5 nmol/L, 9.5 % were never/former smokers with cotinine levels between 5 and 85 nmol/L, 19.4 % had cotinine levels between 85 and 1700 nmol/L and 2.9 % had cotinine levels higher than 1,700 nmol/L at the time of recruitment, compared to 10, 21.8, 6.73, 42.6 and 18.9 % respectively among the cases. Summary information [mean and standard deviation (SD)] for the biomarkers analysed, as well as p for baseline comparison adjusting for smoking exposure, can be found in Table 1.

Several models were estimated considering the four factors and their components, adjusting for age and tobacco exposure. The first model included four factors in which factor 2 and 4 (folate and transsulfuration respectively) were considered as

Table 1 Summary of serum concentrations of b vitamins and metabolites by lung cancer status (891 cases/1,747 controls)

Variable	Mean (SD)		p^1	p^2
	Cases (891)	Controls (1,747)		
Homocysteine (μmol/L)	13.89 (6.79)	13.03 (4.45)	<0.001	0.303
Methionine (μmol/L)	28.57 (7.22)	30.44 (7.57)	<0.001	<0.001
Folate (nmol/L)	15.05 (9.63)	18.25 (15.75)	<0.001	0.001
Vitamin B12 (pmol/L)	380.54 (235.57)	371.12 (215.80)	0.358	0.044
Vitamin B6 (nmol/L)	42.49 (50.74)	52.28 (51.48)	<0.001	0.054
Vitamin B2 (nmol/L)	25.36 (36.47)	28.92 (49.01)	0.054	0.458
Serine (μmol/L)	146.65 (36.26)	148.40 (35.36)	0.303	0.390
Glycine (μmol/L)	299.10 (83.92)	300.67 (89.65)	0.820	0.779
Cysteine (μmol/L)	295.45 (34.75)	299.70 (35.85)	0.001	0.145
Cystathionine (μmol/L)	0.23 (0.15)	0.25 (0.34)	0.188	0.426
Neopterine (nmol/L)	18.41 (6.84)	18.36 (7.09)	0.851	0.218
KTR × 1,000 (10^{-2} nmol/μmol)	22.63 (7.96)	22.60 (7.61)	0.898	0.058

SD standard deviation, *PLP* Pyridoxal 5′-phosphate, *KTR* Kynurenine and Tryptophan ratio, p^1 conditional logistic regression, p^2 conditional logistic regression controlling for smoking exposure, *smoke1* former smoker and cotinine level lower than 5 nmol/L compared to baseline, *smoke2* never/former smoker with cotinine level between 5 and 85 nmol/L compared to baseline, *smoke3* cotinine level between 85 and 1,700 nmol/L compared to baseline, *smoke4* cotinine level higher than 1,700 nmol/L compared to baseline, *baseline* never smoker and cotinine level lower than 5 nmol/L

indirect causes of lung cancer through factor 1 (methionine-homocysteine factor). Factor 2 was also considered as direct cause (Fig. 1). The first model, however, presented very low quality of fit (RMSEA = 0.174, CFI = 0.673, and TLI = 0.510). A second model was hypothesized with the exclusion of factor 4, which improved slightly the goodness-of-fit (RMSEA = 0.161, CFI = 0.717 and TLI = 0.567). In the second model we observed an indirect effect of folate cycle, but its direct effect was not significant ($p = 0.267$) so it was then deleted. The final best-fit model is shown in Fig. 2 and the corresponding results are presented in Table 2. The model was considered adequate according to the following fit indexes: CFI of 0.94, TLI of 0.90, and RMSEA of 0.069.

Three factors included in the final model were found, directly or indirectly, to significantly influence the risk of developing lung cancer at the significance level of 0.05; the effect of each putative factor can be found in Table 2. The standardized weights of each factor component varied from 0.14 to 0.67, being all significantly different from zero ($p < 0.001$).

Factor 1 (methionine-homocysteine metabolism) was composed of methionine and vitamin B12 with positive weights of 0.24 and 0.16 respectively, and homocysteine was excluded to provide a better fit. This means that higher levels of this factor represent higher levels of methylation or homocysteine metabolism. Factor 1 presented a direct and negative effect on lung cancer *probit* score, an increase of one unite in the score representing a change of −0.067 SD in the *probit* score ($p = 0.02$).

Factor 2 (folate cycle) is composed of folate, vitamin B6, vitamin B2 and glycine. Vitamin B6 presents the highest weight (0.63), followed by vitamin B2 (0.42) and folate (0.32) respectively. All weights are positives, so the increases in any of these items provide higher levels in the score of factor 2. Factor 2 showed a positive effect on factor 1 (0.93, $p < 0.001$) which in turn decreased lung cancer risk. In the final model we hypothesized that the effect of factor 2 on lung cancer *probit* was indirect and protective (−0.063, $p = 0.019$).

The first model showed very low goodness-of-fit (see Table 2), and because among the four factors the trans-sulfuration is the one with lack of an established a priori role in lung cancer risk it was excluded.

Finally, factor 3, as expected, showed vitamin B6 as positively associated and neopterin and KTR as negatively associated. Factor 3, representing immune activation, also seemed to be protective (−0.073, $p = 0.021$). Predictably, the direct influence of tobacco exposure on lung cancer risk, independently of the other factors, was the strongest: a change from the lowest category towards the second category of exposure results in an increase of 0.13 SD in the lung cancer *probit* score, as shown in Table 2.

The effects of factor 1, factor 2 and factor 3 presented p values of around 0.02 which would not be significant at the more restricted significant threshold of 0.01, and in this case, could be interpreted just as tendencies.

Despite the fact that tobacco exposure had a significant effect on factor 2 and factor 3, its total indirect effect on lung

Table 2 Results of SEMs for lung cancer *probit* model

	Standardized estimates	*p*	Standardized estimates	*p*
Direct effects of factors on lung cancer				
Homocysteine metabolism (factor 1)	−0.130	0.002	−0.067	0.020
Folate cycle (factor 2)	0.021	0.580		
Immune activation (factor 3)	0.079	0.007	−0.073	0.021
Smoke1	0.132	<0.001	0.129	<0.001
Smoke2	0.112	<0.001	0.111	<0.001
Smoke3	0.451	<0.001	0.448	<0.001
Smoke4	0.428	<0.001	0.430	<0.001
Homocysteine metabolism (factor 1)				
Homocysteine	−0.381	–		
Methionine	0.616	<0.001	0.243	–
Vitamin B12	0.207	<0.001	0.155	<0.001
Folate cycle (factor 2)	0.530	<0.001	0.934	<0.001
Transsulfuration (factor 4)	0.759	<0.001		
Age	−0.37	<0.000	−0.234	0.001
Smoke1	−0.048	0.215	−0.019	0.821
Smoke2	−0.012	0.739	0.170	0.034
Smoke3	0.005	0.906	0.052	0.572
Smoke4	−0.111	0.002	−0.110	0.187
Folate cycle (factor 2)				
Folate	0.392	–	0.323	–
Vitamin B6	0.618	<0.001	0.633	<0.001*
Vitamin B2	0.397	<0.001	0.415	<0.001
Glycine	0.105	<0.001	0.230	<0.001
Age	0.10	0.005	0.102	0.001
Smoke1	−0.04	0.288	−0.050	0.112
Smoke2	−0.07	0.109	−0.108	0.020
Smoke3	−0.28	<0.001	−0.282	<.001
Smoke4	−0.23	<0.001	−0.221	<.001
Immune activation (factor 3)				
Vitamin B6	−0.140	–	0.136	–
Neopterin	0.672	<0.001	−0.672	<0.001
KTR	0.643	<0.001	−0.643	<0.001
Age	0.32	<0.001	−0.319	<0.001
Smoke1	0.105	<0.001	−0.103	0.004
Smoke2	0.059	0.025	−0.058	0.042
Smoke3	−0.048	0.123	0.048	0.137
Smoke4	−0.024	0.381	0.023	0.396
Transsulfuration (factor 4)				
Homocysteine	0.573	–		
Vitamin B6	0.018	0.382		
Cystathionine	0.386	<0.001		
Cysteine	0.723	<0.001		
Age		<0.001		
Smoke1		0.013		
Smoke2		<0.001		

Table 2 continued

	Standardized estimates	*p*	Standardized estimates	*p*
Smoke3		0.702		
Smoke4		0.371		
Indirect effects from factor 2 to lung cancer				
Lung cancer < factor 1 < factor 2	−0.069	0.002	−0.063	0.019
Lung cancer < factor 1< factor 4	−0.098	0.002		
Model fit				
Chi square test	20,731.2 (126 *df*, $p < 0.001$)		10,581.9 (81 *df*, $p < 0.001$)	
RMSEA [90 % Confidence limits]	0.174 [0.171; 0.178]		0.069 [0.065; 0.074]	
CFI	0.673		0.940	
TLI	0.510		0.902	

CFI comparative fit index, *TLI* tucker-lewis index, *RMSEA* root mean square error of approximation, *KTR* kynurenine and tryptophan ratio, *smoke1* former smoker and cotinine level lower than 5 nmol/L compared to baseline, *smoke2* never/former smoker with cotinine level between 5 and 85 nmol/L compared to baseline, *smoke3* cotinine level between 85 and 1700 nmol/L compared to baseline, *smoke4* cotinine level higher than 1,700 nmol/L compared to baseline, *baseline* never smoker and cotinine level lower than 5 nmol/L

cancer through these factors was much lower than its direct effect. The indirect and significant effect of smoking exposure was found for cotinine levels between 85 and 1,700 nmol/L compared to baseline ($\beta = 0.018$, $p = 0.026$), and for cotinine levels higher than 1,700 nmol/L compared to baseline ($\beta = 0.034$, $p = 0.014$) both through factor 2 and factor 1. The indirect effect of former smoker and cotinine level lower than 5 nmol/L, compared to baseline, was found to be significant at 0.05 level ($\beta = 0.008$, $p = 0.008$; baseline: never smoker and cotinine level lower than 5 nmol/L); note that this is the only significant effect which passed the 0.01 threshold.

Discussion

Biomarkers of the one-carbon metabolism (OCM) were recently found to be associated with lung cancer within the EPIC study [8]. In this nested case–control study, we applied SEM to measure four different putative factors (representing mechanisms) of lung carcinogenesis, taking into account tobacco exposure using a combination of blood cotinine levels and smoking status at recruitment. We modelled the risk through four latent factors representing (a) methionine-homocysteine metabolism (factor 1), (b) folate cycle (factor 2), (c) immune activation (factor 3) and (d) transsulfuration (factor 4). Factor 2 was allowed to affect lung cancer indirectly through factor 1. Since Factor 4 has lack of established a priori role in lung cancer risk it was excluded from the initial model and the new hypothesized model presented adequate goodness-of-fit.

Factor 1 showed an inverse association with lung cancer risk; this suggests methionine-homocysteine metabolism may have a putatively protective effect in lung carcinogenesis. Despite the fact that we do not have direct measurements of gene methylation (except in a small subset [31]), we are assuming that the activity level of these reactions is correlated to the total methylation potential of the one-carbon system. Our result is in agreement with previous studies in which global methylation protected against cancer [8, 39, 40]. An earlier study also found that decreased levels of vitamin B12, another essential co-factor in homocysteine metabolism, corresponded with global DNA hypomethylation in squamous cell lung cancer tissue samples [41].

The model with only indirect effect (i.e. no direct effect) of factor 2 on lung cancer risk (through factor 1) was found to have a good model fit, which corroborates the prevailing view that folate mediates its effects on carcinogenesis via its ability to modulate the levels of homocysteine and methionine [13, 42, 43]. Using the same dataset as the current article but applying traditional logistic regression, Johansson et al. [8] also reported an inverse association of folate with lung cancer risk, but this observation was confined to former and current smokers.

In a case–control study, Hartman et al. [5] evaluated the associations between serum vitamins B6, B12, folate and homocysteine with lung cancer. They found vitamin B6 to be inversely associated with lung cancer, which has recently been replicated by Johansson et al. [8] using the same dataset as in the current study.

While the results of Hartman et al. [5] are difficult to interpret due to the limited sample size and absence of never smokers, Johansson et al. [8] observed an association also in never smokers, and described an independent, and equally strong, inverse association of methionine with lung cancer risk. In agreement with our results, Basset et al. [16] suggested that higher intake of vitamin B2 may protect against lung cancer. In our results, vitamin B6 exerted a

protective effect on lung cancer risk via the immune activation (factor 3) as well as indirectly through the folate cycle (factor 2). Hence it is possible that the effect of vitamin B6 noted by Hartman et al. [5], Johansson et al. [8] and here reflects mechanisms independent of the OCM pathway as suggested by Lee et al. [44].

The protective effect of folate cycle may be counter-intuitive considering the pharmacological action of some chemotherapeutic agents (e.g. methotrexate) that are antifolates. Other researchers have proposed that the timing of administration of folate is key to its effect in the different stages of the tumorigenesis process, albeit not in lung cancer [14, 15, 17].

There has been a small number of randomised clinical trials where the effects of folate and certain B-vitamins (B6 and B12) supplementation on cancer risk were investigated, raising concerns about safety of folic acid fortification. The studies have been conducted in the US where fortification was implemented early. However, these studies reported conflicting results, with some authors citing the differentials in dosage and baseline levels of folate as possible explanations [39, 40]. The results from the current analysis suggest that the compounds within the OCM could interact in a more dynamic manner than previously thought.

We have found an inverse association between lung cancer risk and factor 3, which represents immune activation and inflammation. Hence higher levels of neopterin or KTR produced a higher *probit* score (higher risk), whereas the inverse applies to vitamin B6. High levels of vitamin B6 have been associated with reduction in inflammation and oxidative stress [41], as well as macrophage-related immune response [42, 43].

To our knowledge, this is the first study which has attempted to evaluate the roles and the relative contributions of the multiple components of the OCM, revealing indirect effects, in relation to lung cancer. The study had a prospective design, in which biomarkers in blood were collected several years before disease onset. The main advantage of SEM is the possibility to account for multiple, complex hypothesized causal pathways, and also of dealing rigorously with collinearity. In contrast to usual multiple regression models, which account for the relative contribution of several variables to the one observed effect, in which strong collinearity can be problematic, in SEM the factors are based on correlations, thus avoiding such limitation.

There are also limitations to the current study: blood samples were drawn on one occasion only for each participant, thus preventing us from investigating time-related relations among the investigated factors. The biochemical pathways that we attempted to model are very complex, and it was not possible to include all biomarkers potentially involved due to a lack of complete information. We did not include the effects of polymorphisms related to OCM. We have chosen to use biomarkers of OCM components (e.g. B-vitamins) measured in the blood rather than estimations from food frequency questionnaires. While confounding from dietary intakes are still possible, this is considered outweighed by the advantages of the increased accuracy and biological relevance of using biomarkers measurements. The large number of statistical tests performed in the final model could inflate type I error and bias the reported results. However, many of the tests were very robust ($p < 0.001$) and SEM analysis is usually more robust compared to regression since it considers measurement error [45]. In the structural model three effects presented p around 0.02, and their interpretations require careful consideration. These values may present less efficiency and the tendencies of the effects need to be confirmed by other investigators and in other populations.

Also, since tobacco exposure is thought to modulate the levels of B vitamins, homocysteine and certain other OCM components [13], this would mean that tobacco acts as a classic confounder in the SEM. Despite our attempt to adjust for its effect by combining self-reported smoking status with blood cotinine levels, it is still possible that residual confounding exists.

Conclusion

In conclusion, our results support the roles of methionine-homocysteine metabolism which is indicative of the DNA methylation potential (including indirect effects of folate cycle), and immune activation in lung cancer risk. Regarding the involvement of components in the OCM, the folate cycle and immune system in lung cancer carcinogenesis, it may be important to consider complex pathways rather than the activities of single vitamins or nutrients. Hence the use of structural equation modelling may be key in dealing with such complex theoretical models where the degree of collinearity is high. In the model presented in the current study, tobacco smoking remains the exposure with the strongest impact on lung cancer risk.

Acknowledgments This paper was supported by the World Cancer Research Fund (grant number 2007/16-UK). This paper forms part of VTB's Ph.D. Dissertation. VTB was supported by the National Council for Scientific, Technological Development (CNPq) and Capes Foundation, Brazil.

Conflict of interests The authors declare that they have no conflict of interest.

References

1. World Health Organization. In: Cancer. Fact sheet N°297. 2012. http://www.who.int/mediacentre/factsheets/fs297/en. Accessed 25 Aug 2012.

2. Doll R, Hill AB. Smoking and carcinoma of the lung: preliminary report. Br Med J. 1950;2:739–48.
3. Ames BN. Low micronutrient intake may accelerate the degenerative diseases of aging through allocation of scarce micronutrients by triage. PNAS. 2006;103(47):17589–94.
4. Hartman TJ, Woodson K, Stolzenberg-Solomon R, Virtamo J, Selhub J, Barrett MJ, Albanes D. Association of the b-vitamins pyridoxal 5′-phosphate (b6), B12, and folate with lung cancer risk in older men. Am J Epidemiol. 2001;153(7):688–94.
5. Johansson M, Van Guelpen B, Vollset SE, Hultdin J, Bergh A, Key T, Midttun O, Hallmans G, Ueland PM, Stattin P. One-carbon metabolism and prostate cancer risk: prospective investigation of seven circulating b vitamins and metabolites. Cancer Epidemiol Biomarkers Prev. 2009;18(5):1538–43.
6. Ulrich CM, Reed MC, Hijhout HF. Modeling folate, one-carbon metabolism, and DNA methylation. Nutr Rev. 2008;66(1):S27–30.
7. Ulrich CM, Neuhouser M, Liu AY, Boynton A, Gregory JF 3rd, Shane B, James SJ, Reed MC, Nijhout HF. Mathematical modelling of folate metabolism: predicted effects of genetic polymorphisms on mechanisms and biomarkers relevant to carcinogenesis. Cancer Epidemiol Biomarkers Prev. 2008;17(7):1822–31.
8. Johansson M, Relton C, Ueland PM, Vollset SE, Midttun Ø, Nygård O, Slimani N, Boffetta P, Jenab M, Clavel-Chapelon F, Boutron-Ruault MC, Fagherazzi G, Kaaks R, Rohrmann S, Boeing H, Weikert C, Bueno-de-Mesquita HB, Ros MM, van Gils CH, Peeters PH, Agudo A, Barricarte A, Navarro C, Rodríguez L, Sánchez MJ, Larrañaga N, Khaw KT, Wareham N, Allen NE, Crowe F, Gallo V, Norat T, Krogh V, Masala G, Panico S, Sacerdote C, Tumino R, Trichopoulou A, Lagiou P, Trichopoulos D, Rasmuson T, Hallmans G, Riboli E, Vineis P, Brennan P. Serum B vitamin levels and risk of lung cancer. JAMA. 2010;303(23):2377–85.
9. Robertson CE, Vineis P. Biomarkers for nutrient-gene interactions. In: Choi S-W, Friso S, editors. Nutrient-gene interactions in cancer. Boca Raton: CRC Press; 2006. p. 37–55.
10. Choi S-W, Friso S. Interaction between folate and methylenetetrahydrofolate reductase gene in cancer. In: Choi S-W, Friso S, editors. Nutrient-gene interactions in cancer. Boca Raton: CRC Press; 2006. p. 57–85.
11. Lee MS, Asomaning K, Su L, Wain JC, Mark EJ, Christiani DC. MTHFR polymorphisms, folate intake and carcinogen DNA adducts in the lung. Int J Cancer. 2012;131:1203–9.
12. Das PM, Signal R. DNA methylation and cancer. J Clin Oncol. 2004;22:4632–42.
13. Selhub J. Homocysteine metabolism. Annu Rev Nutr. 1999;19:217–46.
14. Ulrich CM, Potter JD. Folate and cancer—timing is everything. JAMA. 2007;297(21):2408–9.
15. Kim YI. Folate: a magic bullet or a double edged sword for colorectal cancer prevention? Gut. 2006;55(10):1387–9.
16. Bassett JK, Hodge AM, English DR, Baglietto L, Hopper JL, Giles GG, Severi G. Dietary intake of B vitamins and methionine and risk of lung cancer. Eur J of Clin Nutr. 2012;66(2):182–7.
17. Ulrich CM, Potter JD. Folate supplementation: too much of a good thing? Cancer Epidemiol Biomarkers Prev. 2006;15(2):189–93.
18. Leklen JE. Vitamin B6. In: Ziegler EE, Filer Jr LJ, editors. Present Knowledge in nutrition. Washington: ILSI Press; 1996. p. 174–83.
19. Davis SR, Quinlivan EP, Stacpoole PW, Gregory JF 3rd. Plasma glutathionine and cystathionine concentrations are elevated but cysteine flux is unchanged by dietary vitamin B-6 restriction in young men and women. J Nutr. 2006;136:373–8.
20. Price JM, Brown RR, Yess N. Testing the functional capacity of the tryptophan-niacin pathway in man by analysis of urinary metabolites. Adv Metab disord. 1965;2:159–225.
21. Murray MF. Tryptophan depletion and HIV infection: a metabolic link to pathogenesis. Lancet Infect Dis. 2003;3:644–52.
22. Boasso A, Vaccari M, Hryniewics A, Fuchs D, Nacsa J, Cecchinato V, Andersson J, Franchini G, Shearer GM, Chougnet C. Regulatory T-cell markers, indoleamine 2,3-dioxygenase, and virus levels in spleen and gut during progressive simian immunodeficiency virus infection. J Virol. 2007;81:11593–603.
23. Moreau M, Lesgate J, Verrier D, Mormed C, Kelly KW, Castanon N. Bacille Calmette-Guérin inoculation induces chronic activation of peripheral and brain indoleamine 2,3-dioxygenase in mice. J Infect Dis. 2005;192:537–44.
24. Silva NM, Rodrigues CV, Santoro MM, Reis LF, Alvarez-Leite JI, Gazzinelli RT. Expression of indoleamine 2,3-dioxygenase, tryptophan degradation, and kynurenine formation during in vivo infection with Toxoplasma gondii: induction by endogenous gamma interferon and requirement of interferon regulatory factor 1. Infect Immun. 2002;70:859–68.
25. Melchior D, Sève B. Le Floc'h N. Chronic lung inflammation affects plasma amino acid concentration in pigs. J Anim Sci. 2004;82:1091–9.
26. Murr C, Widner B, Wirleitner B, Fuchs D. Neopterin as a marker for immune system activation. Curr Drug Metab. 2002;3(2):175–87.
27. Hoffmann G, Wirleitner B, Fuchs D. Potential role of immune system activation-associated production of neopterin derivatives in humans. Inflamm Res. 2003;52(8):313–21.
28. Sucher R, Schroecksnadel K, Weiss G, Margreiter R, Fuchs D, Brandacher G. Neopterin, a prognostic marker in human malignancies. Cancer Lett. 2010;287(1):13–22.
29. Vaissière T, Cuenin C, Paliwal A, Vineis P, Hoek G, Krzyzanowski M, Airoldi L, Dunning A, Garte S, Hainaut P, Malaveille C, Overvad K, Clavel-Chapelon F, Linseisen J, Boeing H, Trichopoulou A, Trichopoulos D, Kaladidi A, Pall D, Krogh V, Tumino R, Panico S, Bueno-De-Mesquita HB, Peeters PH, Kumle M, Gonzalez CA, Martinez C, Dorronsoro M, Barricarte A, Navarro C, Quiros JR, Berglund G, Janzon L, Jarvholm B, Day NE, Key TJ, Saracci R, Kaaks R, Riboli E, Hainaut P, Herceg Z. Quantitative analysis of DNA methylation after whole bisulfitome amplification of a minute amount of DNA from body fluids. Epigenetics. 2009;4(4):221–30.
30. Paliwal A, Vaissière T, Herceg Z. Quantitative detection of DNA methylation states in minute amounts of DNA from body fluids. Methods. 2010;52(3):242–7.
31. Vineis P, Chuang SC, Vaissière T, Cuenin C, Ricceri F. Genair-EPIC Collaborators, Johansson M, Ueland P, Brennan P, Herceg, Z. DNA Methylation changes associated with cancer risk factors and blood levels of vitamin metabolites in a prospective study. Epigenetics. 2011;6(2):195–201.
32. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondière UR, Hémon B, Casagrande C, Vignat J, Overvad K, Tjønneland A, Clavel-Chapelon F, Thiébaut A, Wahrendorf J, Boeing H, Trichopoulos D, Trichopoulou A, Vineis P, Palli D, Bueno-De-Mesquita HB, Peeters PH, Lund E, Engeset D, González CA, Barricarte A, Berglund G, Hallmans G, Day NE, Key TJ, Kaaks R, Saracci R. European prospective investigation cancer and nutrition (EPIC): study populations and data collection. Public Health Nutr. 2002;5:1113–24.
33. McCutcheon AL. Latent class analysis. No. 07-064. University Paper Series on Quantitative Applications in the Social Sciences. Newbury Park, CA: Sage, 1987.
34. Skrondal A, Rabe-Hesketh S. Generalized latent variable modeling: multilevel, longitudinal, and structural equation models. Washington: Chapman & Hill/CRC; 2004.

35. Baltar VT, Xun WW, Chuang SC, Relton C, Ueland PM, Vollset SE, Midttun Ø, Johansson M, Slimani N, Jenab M, Clavel-Chapelon F, Boutron-Ruault MC, Fagherazzi G, Kaaks R, Rohrmann S, Boeing H, Weikert C, Bueno-de-Mesquita HB, Boshuizen HC, van Gils CH, Peeters PH, Agudo A, Barricarte A, Navarro C, Rodríguez L, Castaño JM, Larrañaga N, Pérez MJ, Khaw KT, Wareham N, Allen NE, Crowe F, Gallo V, Norat T, Tagliabue G, Masala G, Panico S, Sacerdote C, Tumino R, Trichopoulou A, Lagiou P, Bamia C, Rasmuson T, Hallmans G, Roswall N, Tjønneland A, Riboli E, Brennan P, Vineis P. Smoking, secondhand smoke, and cotinine levels in a subset of EPIC cohort. Epidemiol Biomarkers Prev. 2011;20(5):869–75.
36. Byrne BM. Structural equation modeling with Mplus: Basic concepts, applications, and programming (Multivariate Applications Series). New York: Routledge; 2011.
37. Muthén LK, Muthén BO. Mplus User's Guide. 6th ed. Los Angeles: Muthén & Muthén; 2010.
38. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods. 1996;1996(1):130–49.
39. Baluz K, do Carmo MGT, Rosas G. The role of folic acid on oncologic prevention and intervention: review. Revista Brasileira de Cancerologia. 2002;48(4):597–607.
40. Kim YI. Role of folate in colon cancer development and progression. J Nutr. 2003;133:3731S–9S.
41. Piyathilake CJ, Johanning GL, Macaluso M, Whiteside M, Oelschlager DK, Heimburger DC, Grizzle WE. Localized folate and vitamin B-12 deficiency in squamous cell lung cancer is associated with global DNA hypomethylation. Nutr Cancer. 2000;37(1):99–107.
42. Selhub J. The Many Facets of Hyperhomocysteinemia: studies from the Framingham Cohorts. J Nutr. 2006;136(6):1726S–30S.
43. Scott JM. Folate and vitamin B12. Proc Nutr Soc. 1999;58(2):441–8.
44. Lee JE, Li H, Giovannucci E, Lee IM, Selhub J, Stampfer M, Ma J. Prospective study of plasma vitamin B6 and risk of colorectal cancer in men. Cancer Epidemiol Biomarkers Prev. 2009;18(4):1197–202.
45. Brunner J, Austin PC. Inflation of type I error rate in multiple regression when independent variables are measured with error. Can J Stat. 2009;37(1):33–46.