Plasma vitamin C concentrations and risk of incident respiratory diseases and mortality in the European Prospective Investigation into Cancer-Norfolk population-based cohort study

Running Head: Vitamin C levels and incident respiratory diseases

Phyo Kyaw Myint, Andrew M Wilson, Allan B Clark, Robert N Luben, Nicholas J Wareham, Kay-Tee Khaw

School of Medicine, Medical Sciences & Nutrition, University of Aberdeen, Aberdeen, UK (PKM); Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, UK (PKM, AMW, ABC); Clinical Gerontology Unit, Department of Public Health and Primary Care, School of Clinical Medicine, University of Cambridge, Cambridge, UK (PKM, RNL, KTK); MRC Epidemiology Unit, Cambridge, UK (NJW)

Professor Phyo Kyaw **MYINT;** *Professor of Medicine of Old Age*
phyo.myint@abdn.ac.uk

Professor Andrew M **WILSON;** *Professor of Respiratory Medicine*
a.m.wilson@uea.ac.uk

Dr Allan B **CLARK;** *Senior Lecturer in Medical Statistics*
Allan.clark@uea.ac.uk

Mr Robert N **LUBEN;** *Senior Research Associate*
Robert.luben@srl.cam.ac.uk

Professor Nicholas J **WAREHAM;** *Director of MRC Epidemiology Unit*
Nick.wareham@mrc-epid.cam.ac.uk

Professor Kay-Tee **KHAW;** *Professor of Clinical Gerontology*
Kk101@medschl.cam.ac.uk

Corresponding author:
Phyo Kyaw Myint
Room 4:013Polwarth Building,
School of Medicine, Medical Sciences and Nutrition
University of Aberdeen, Foresterhill,
Aberdeen
AB25 2ZD
Tel:+44 (0) 1224 437957 | Fax: +44 (0) 1224 437911
Mail to: phyo.myint@abdn.ac.uk

Sources of support: The EPIC-Norfolk study was supported by grants from the Medical Research Council and Cancer Research UK.

Abbreviations:

COPD: Chronic Obstructive Pulmonary Disease

CV: Coefficient of variation

ENCORE: East Norfolk Commission Record

EPIC-Norfolk: European Prospective Investigation into Cancer-Norfolk

FEV1: Forced expiratory volume in first second

FFQ: Food frequency questionnaire

FVC: Forced vital capacity

ICD: International classification of disease

TDI: Townsend deprivation index.

Abstract

Background: Cancerous and non-cancerous respiratory diseases are common and contribute significantly to global disease burden. We aim to quantify the association between plasma vitamin C concentrations as an indicator of high fruit and vegetable consumption and the risk of incident respiratory diseases and associated mortality in a general population.

Methods: 19,357 men and women aged 40-79 years without prevalent respiratory diseases at the baseline (1993-1997) and participating in the European Prospective Investigation into Cancer (EPIC)-Norfolk study in the UK were followed through March 2015 for both incidence and mortality from respiratory diseases.

Results: There were a total of 3914 incident events and 407 deaths due to any respiratory diseases (excluding lung cancers), 367 incident lung cancers and 280 lung cancer deaths during the follow up (total person years >300,000 years). Cox proportional hazards models showed, persons in the top quartiles of baseline plasma vitamin C concentrations had a 43% lower risk of lung cancer (HR 0.57; 95%CI:0.41-0.81) than did those in the bottom quartile, independently of potential confounders. The results are similar for non-cancerous any respiratory disease (HR 0.85;0.77-0.95), chronic respiratory diseases (HR0.81;0.69-0.96), and pneumonia (HR0.70;0.59-0.83). The corresponding values for mortality were 0.54(0.35-0.81), 0.81(0.59-1.12), 0.85(0.44-1.66), and 0.61(0.37-1.01), respectively. Confining analyses to non-smokers showed 42% and 53% risk reduction of non-smoking related lung cancers incidence and death.

Conclusions: Higher levels of vitamin C concentrations as a marker of high fruit and vegetable consumption reduces the risk of cancerous and non-cancerous respiratory illnesses including non-smoking related cancers incidence and deaths.

Keywords

- Fruit and vegetable consumption
- Vitamin C
- Lung cancer
- Chronic respiratory disease
- Pneumonia
- Incidence
- Mortality

Introduction

Respiratory diseases are major causes of mortality and morbidity worldwide. In the UK alone, respiratory disease costs the NHS and society £6.6 billion: £3 billion in costs to the care system, £1.9 billion in mortality costs and £1.7 billion in illness costs [1]. Irreversible progressive chronic respiratory diseases and pneumonia are common in middle and older age and contribute significantly to this burden. Whilst lung cancer incidence is declining in some countries which imposed smoking ban, there remains risk of lung cancers in general as well as the risk of non-smoking related lung cancer.

Whilst the intake of high fruit and vegetables depicted by plasma vitamin C concentrations are linked to mortality [2] and chronic cardiovascular conditions including stroke [3-7]. Whether such dietary health behaviour is linked to respiratory diseases including pneumonia is less well understood. Plasma vitamin C levels can be regarded as a reasonably reliable measure of not just the intake of fruit and vegetables, but the in vivo biologically available vitamin C for human physiological functions because it cannot be synthesized in the body. It is because the main source of vitamin C is from dietary fruit, vegetables, and plant foods in humans but is easily denatured by food preparation methods (e.g. via cooking in water, roasting, or grilling) [8]. It also has a short half-life (≈30 min) in the blood [9,10]. Therefore, random plasma vitamin C level is most likely to reflect an individual's habitual dietary pattern as well as method of food preparation.

The relationship between plasma vitamin C concentrations and the incidence of chronic obstructive pulmonary disease (COPD), pulmonary fibrosis, pneumonia and respiratory mortality including lung cancer deaths, in particular non-smoking related lung cancers and

59 their associated mortality risks were poor understood. We explored these relationships in the
60 European Prospective Investigation into Cancer (EPIC)–Norfolk. Our secondary objective
61 was to examine whether the relationship, if existed, vary by age, sex, smoking status,
62 physical activity level, and lung function assessed by forced expiratory volume in the first
63 second (FEV1). We then further assessed the relationship between higher level of fruit and
64 vegetable consumption and non-smoking related incidence and mortality for these conditions.
65
66

Subjects and Methods

Participants

Participants were drawn from the EPIC-Norfolk prospective population study. The detailed recruitment method and study protocol of EPIC-Norfolk were described previously [11]. Briefly, all eligible community-dwelling adults (aged 40-79 years at the baseline (1993-1997)) from 35 participating general practices in Norfolk, UK, were invited to participate. A total of 30,445 (~40% response) persons provided written consent to participate in the study (99.6% British Caucasians). The Norwich Local Research Ethics Committee approved the study.

Measurements

At the time of the baseline survey, participants completed a detailed health and lifestyle questionnaire and attended health check clinic where trained nurses performed clinical assessments. The data collection methods for variables included in this study are described details in the online supplement [12-17].

Assessment of Vitamin C levels

After overnight storage in a dark box at 4–7 °C, non-fasting blood samples were centrifuged at 2100×*g* for 15 min at 4 °C. Plasma vitamin C was measured from blood collected one year later when funding was available using fluorometric assay within 1 week of sampling [18]. The mean coefficients of variation (CV) were 33.2 μmol/L at the lower end and 102.3 μmol/L at the upper end.

Outcome ascertainment

Incident cases were ascertained by using death certificate data and hospital record linkage based on UK Office of National Statistics using International Classification of Disease (ICD), revisions 9 and 10 and National Health Service hospital information systems through ENCORE ((ENCORE – **E**ast **N**orfolk **CO**mmission **RE**cord). Accuracy of this method has been previously validated [19]. Hospitalisation record was linked up to end of March 2015 and death was monitored until end March 2015. Conditions of interest are any respiratory conditions (ICD codes 460-519; J00-J99, chronic respiratory conditions ICD codes 490-496; J40-47,pneumonia, ICD 10 J12-J18, B012,B052,B953,B960,B961,J100, J110,J851, Lung cancers (C33-C34) and asthma (J45-J46). Whilst incident and morality from asthma was included in all non-cancerous respiratory diseases, it was specifically excluded from outcomes as adult/older age onset asthma may have other mechanistic pathways which are non-dietary.

Statistical analysis

Statistical analysis was performed using Stata Version 14.1/SE (StatCorLP, USA. Texas, USA). All tests are two-sided and, as the study is exploratory, no adjustment for multiple comparisons has been made. Any participants with missing value for any of the variables included in the analyses were excluded in individual analyses. Plasma vitamin C concentration was split into quartiles, with cut-points 41, 54, and 66 μmol/L. The percentage of predicted normal were calculated for FVC and FEV1 (FVC% predicted and FEV1% predicted respectively) based on the Steel and Coal formula [20]. The incidence of respiratory conditions and respiratory related mortality are reported descriptive, firstly overall and then stratified by the quartiles of plasma vitamin C concentrations.

Cox proportional hazards models were constructed to determine the association between plasma vitamin C concentrations and the subsequent risk of all respiratory conditions, chronic respiratory disease, and pneumonia. The risks of mortality for each of the selected outcomes were also assessed. To estimate the independent contribution of vitamin C concentration we controlled for 1) age and sex; 2) age, sex, and respiratory function assessed using FVC and FEV1; 3) age, sex, respiratory function, and lifestyle factors (smoking, alcohol consumption, physical activity and BMI); 4) age, sex, respiratory function, lifestyle factors, occupational social class and Townsend index of deprivation; 5) as in model 3 with additional adjustment for prevalent diabetes, myocardial infarction (MI) and stroke; 6) as in model 4 with additional adjustment for history of diabetes, MI and stroke; and finally model 7 was constructed as in model 6 after excluding participants who were taking vitamin C containing supplements.

To assess the differences in the relationship between plasma concentration of vitamin C at the baseline and the subsequent incidence of respiratory diseases including mortality specifically in different age groups (younger and older (<65 years and >=65 years)), sex (men and women), smoking status (current smokers and non/ex-smokers), physical activity levels (inactive and active), and respiratory function (good and poor function; people with good function defined as people in the top quartile of FEV1 values) the stratified analyses were repeated for model 6.

Analyses were then repeated for main analyses for final two models (models 6 and 7) confining first to non-current smokers and then to never-smokers.

142

Results

A total of 25,639 individuals attended their first health check in EPIC-Norfolk during 1993-1997, of these 1,410 were excluded as they had an existing diagnosis of cancer leaving 24,229 individuals. Of these, 2205 had self-reported respiratory diseases and/or taking bronchodilators or drugs for asthma or COPD at the baseline; this leaves 22,024 individuals. After exclusion of 2,685 individuals who did not have vitamin C measurements and further 3 participants with missing follow-up data we included a total of 19,336 men and women in the current study**.** The mean follow-up was 16.54 years. The total follow-up period for the incidence was 319,937 person-years and for the mortality outcome was 336,473 person-years. There were a total of 3,914 incident cases and 407 deaths during the respective follow up periods.

Table 1 shows the sample characteristics by quartiles of plasma vitamin C concentrations. The quartile values were =< 41 µmol/L, 42-54 µmol/L, >54-=<66.0 µmol/L, and >66.0 µmol/L for quartiles 1,2, 3 and 4, respectively. This shows significant trends for all selected variables across the plasma vitamin C categories. People with the higher levels of plasma vitamin C levels were younger, had better baseline lung functions, materially less likely to be deprived, lower proportion of men, less likely to be current smokers and consumed less amount of alcohol, had higher level of education and occupational social class, physically more active, and had lower prevalence of chronic co-morbid conditions. The crude rates for all the selected outcomes suggest a direct relationship between higher baseline plasma vitamin C concentrations and a better outcome (lower event rates).

Table 2 shows the hazards ratios and their corresponding 95% confidence intervals for development of incident any respiratory diseases including chronic respiratory diseases (e.g. COPD/pulmonary fibrosis etc.) and pneumonia over the follow up period. Compared to the people in the bottom quartiles, the people in the highest levels of baseline plasma ascorbic concentrations had significant 15%, 19% 30% and 47% relative risk reduction of having any respiratory conditions, chronic respiratory diseases, hospitalised pneumonia, and lung cancer respectively in adjusted model (model 6). The corresponding relative risk reduction for mortality from these conditions were 13%, 37% and 34% and 52%, respectively, although these did not reach to statistical significance (**Table 3**) with the exception of lung cancer, which showed relative risk reduction of 46%. Excluding those participants who took vitamin C containing supplements did not alter the association (model 7, Tables 2 &3).

Figure 1 (a-d) show the stratified analyses by age groups (younger and older (<65 years and >=65 years)), sex (men and women), smoking status (current smokers, non/ex-smokers), physical activity (active, inactive), and respiratory function (good function defined as top quartile of FEV1 value, and poor function defined as the lower three quartiles) for all outcomes examined for those in top quartile of vitamin C concentrations compared to the bottom quartile. The benefit of high levels of fruit and vegetable consumption depicted by random plasma vitamin C concentrations was associated with younger age but generally all subgroups showed consistent trends towards benefit with regard to all outcomes examined. People with lower FEV1 also appear to be benefited from higher fruit and vegetable consumption compared to people with better FEV1. With reduced statistical power most of the sub-analyses shown in **Supplementary Table 1** for mortality outcome, whilst most of the estimates did not show any significant benefit with regard to mortality due to these conditions

but the point estimates are in consistent with potential benefit across age group, sex, smoking status, physical activity, and the level of FEV1.

Supplementary Table 2 shows the results confining to non-smokers and never-smokers for both incidence and mortality of all these respiratory diseases. With lower numbers whilst the 95%CI are wider, the point estimates show similar direction. Of note, the non-smoking related lung cancer incidence and mortality showed stronger benefit 42% and 53% relative risk reduction in non-current smokers in people with top quartiles of vitamin C concentrations compared to the bottom quartile. **Supplementary Figure 1** shows the Kaplan-Meier estimates of survival based on different events. The graphs are top-left for all respiratory events, top right for lung cancer events, bottom-left for chronic respiratory disease event, and bottom right for pneumonia events.

Discussion

Plasma vitamin C concentration is a good biomarker of plant food, namely fruit and vegetables intake. We found for the first time that the higher habitual intake of fruit and vegetables depicted by higher plasma vitamin C concentrations at the baseline predict future incidence of any respiratory diseases (excluding cancer), chronic respiratory diseases and hospitalised pneumonia in a general population. Further we demonstrate significant relative risk reduction in lung cancer incidence as well as lung cancer related mortality by higher level of baseline plasma vitamin C concentrations. The observed effects are even more pronounced for non-current smokers. The benefit observed with each higher quartiles of plasma vitamin C is fairly consistent and the most beneficial relative risk reduction was observed in the top quartile and the cut off point for the highest quartile in this study was >66 µmol/L. Given that one serving of fruit and vegetables in our cohort equates to 20-µmol or 1-SD increase in plasma vitamin C concentration [2], it can be extrapolated from our results that at the level of 3.5 portions of fruit and vegetable consumption could lead to significant health benefit within a general population.

It is well recognised that fresh rather than cooked or boiled fruit and vegetables are a richer source of vitamin C. Due to their antioxidant property, higher levels of vitamin C are thought to be associated with reduced inflammatory process/insult. However, there appears to be distinct lack of nutritional interventions in respiratory diseases except vitamin supplementation with only a few short term studies evaluating supplementation with vitamins. The link between vitamin C intake and lung cancer incidence has been previously investigated [21]. We have furthered this knowledge and shown that high consumption of

fruit and vegetables also potentially have substantial survival benefit with relative risk reduction of 46%.

One short term study of Vitamin C and E showed no change in lung function in 24 COPD patients [22]. In a larger 5-year trial of more than 20,000 patients with coronary disease, other occlusive arterial disease, or diabetes, there was no difference in lung function or COPD related outcomes [23]. A Cochrane review of vitamin C treatment for pneumonia identified 3 prophylactic studies conducted more than 30 years ago involving 2335 people from which 37 developed community-acquired pneumonia. In all of these studies there was a significant reduction in the development of pneumonia however only one study (of 674 marine recruits) was a randomised trial and the evidence was felt to be too weak to advocate prophylactic use of vitamin C to prevent pneumonia in the general population [24].

Data from the Third National Health and Nutrition Examination Survey, a US population-based cross-sectional study suggest that serum vitamin C concentration is independently related to lung function as assessed by FEV1 [25] whereas serum vitamin C concentration is associated with FVC in patients with airflow obstruction [26]. Other authors have shown the concentration of vitamin C to be inversely related to all-cause mortality in adults with obstructive lung disease [27] and mean intake of vitamin C is significantly lower in patients with COPD than controls [28].

Reactive oxygen species are produced by inflammatory cells including neutrophils and macrophages or are inhaled as environmental agents or cigarette smoke. Oxidants result in cellular damage and may result in apoptosis or cellular necrosis, induce mucous secretion or alter remodelling of extracellular matrix [29] all of which may contribute to the

253 pathophysiology of obstructive lung diseases. Indeed, vitamin C has recently be shown to
254 prevent smoke-induced emphysema in mice that cannot synthesise vitamin C suggesting that
255 vitamin C reduced the oxidative stress caused by cigarette smoking [30] and therefore are not
256 simply a reflection of diet. Interestingly, in the current study, people in the lowest quartile of
257 vitamin C had the highest proportion of those with smoking history – Vitamin C has been
258 shown to be required to have an adequate immune response to influenza virus infection in
259 mice [31] and micronutrient deficiencies (including vitamin C) are associated with impaired
260 immune response and higher burden of respiratory infections in elderly humans [32].
261
262 There are some limitations which worth discussing. The initial response rate was modest
263 (~40%). Nevertheless, the baseline characteristics of the study population were similar to
264 those of other UK population samples, with the exception of slightly lower prevalence of
265 smokers [11] which may be related to healthy responder bias. Whilst the lower prevalence of
266 smoking may have some impact on vitamin C levels, the internal relationships between
267 vitamin C levels and incidence respiratory diseases are unlikely to be affected by this.
268 Moreover, the truncation of distribution is more likely to attenuate the associations. .
269 Although follow-up is virtually complete, using hospital record and death certificate linkage
270 approach may underestimate events that are not admitted to the hospital e.g. milder forms of
271 pneumonia.
272
273 The use of self-reported respiratory conditions may have missed some prevalent respiratory
274 conditions. Plasma vitamin C and other covariates, were made at baseline. These measures as
275 well as lifestyle behaviours, which may affect vitamin C concentration, may have changed
276 over the follow-up period. Due to missing data in some of the variables included in the
277 models, the number of participants included in the analysis reduced with higher level of

adjustments but the findings were consistent and adjustments did not particularly attenuate the results.

We addressed confounding and reverse causality issues by comprehensive adjustment of possible confounders. We also performed sensitivity analysis for supplements usage. The incident cases were identified after 1998, at least one year after the baseline. Prospective relationship between baseline vitamin C concentrations and incidence and mortality also suggests potential causal link, or at least the marker of the conditions examined in this study. We performed sensitivity analyses by confining analyses to those who were non-current smokers and never-smokers.

To conclude, we found that high plasma vitamin C concentrations at baseline are significantly associated with the lower incidence of lung cancer, any respiratory illness other than lung cancer, chronic respiratory diseases and pneumonia. Additionally, those in highest consumption of fruit and vegetables (Quartile 4 of plasma vitamin C concentrations) also had 46% relative risk reduction from lung cancer mortality (53% in those who were non-current smokers at the baseline) compared to the bottom quartile. Therefore, the plasma vitamin C level may be useful in identification of those at risk of these both cancerous and non-cancerous respiratory diseases at the population level. As vitamin C level in plasma is associated with higher fruit and vegetable consumption, increased consumption of fruit and vegetables may reduce the risk of these respiratory conditions and associated health care cost globally.

300 **Figure Legends**

301 **Figure 1:** Hazard ratios for study outcomes in people in the highest quartile group (Q4)
302 compared to the lowest quartile group (Q1) stratified by important factors for model 6.

303 **Figure 1 (a):** Incident respiratory illness

304 **Figure 1 (b):** Incident lung cancer

305 **Figure 1 (c):** Incident chronic respiratory diseases

306 **Figure 1 (d):** Incident pneumonia

307

Acknowledgements

The authors would like to thank the participants of the EPIC-Norfolk cohort. We thank the nutritionist team and data management team of the EPIC-Norfolk cohort. The EPIC-Norfolk study was supported by grants from the Medical Research Council and Cancer Research UK. Funders and sponsors had no role in design and the data collection, analysis, and interpretation of data and the writing of the article and the decision to submit it for publication.

Contributors

PKM: design, write up, primary responsibility for final content

AMW: design, write up

ABC: data analysis, write up

RNL: data linkage, write up

NJW: design, provide data, write up

KTK: design, provide data, write up

Conflict of interest

None for all authors.

Supplementary information is available at EJCN's website.

References

[1] Gupta R and Limb E. The Burden of Lung Disease: A statistics report from the British Thoracic Society. 2nd edition. 28 June 2006 [accessed 27 Nov 2014]

[2] Khaw KT, Bingham S, Welch A, Luben R, Wareham N, Oakes S, et al. Relation between plasma ascorbic acid and mortality in men and women in EPIC-Norfolk prospective study: a prospective population study. European Prospective Investigation into Cancer and Nutrition. *Lancet*. 2001;357:657-63.

[3] Myint PK, Luben RN, Welch AA, Bingham SA, Wareham NJ, Khaw KT. Plasma vitamin C concentrations predict risk of incident stroke over 10 y in 20 649 participants of the European Prospective Investigation into Cancer Norfolk prospective population study. *Am J Clin Nutr.* 2008;87:64-9.

[4] Myint PK, Luben RN, Wareham NJ, Bingham SA, Khaw KT. Combined effect of health behaviours and risk of first ever stroke in 20,040 men and women over 11 years' follow-up in Norfolk cohort of European Prospective Investigation of Cancer (EPIC Norfolk): prospective population study. *BMJ.* 2009;338:b349

[5] Ness AR, Powles JW, Khaw KT. Vitamin C and cardiovascular disease: a systematic review. *J Cardiovasc Risk* 1996;3:513-21.

[6] Yokoyama T, Date C, Kokubo Y, Yoshiike N, Matsumura Y, Tanaka H. Serum vitamin C concentration was inversely associated with subsequent 20-year incidence of stroke in a Japanese rural community. The Shibata Study. *Stroke* 2000;31:2287-94.

[7] Voko Z, Hollander M, Hofman A, Koudstaal PJ, Breteler MM. Dietary antioxidants and the risk of ischemic stroke: the Rotterdam Study. *Neurology* 2003;61:1273-5.

[8] Food Standard Agency. EVM review of vitamin C. Eat well, be well. Helping you make healthier choices. Internet: http://www.food.gov.uk/multimedia/pdfs/evm_c.pdf (accessed 15 September 2014).

[9] Hickey S, Roberts H. Ascorbate: the science of vitamin C. Napa, CA: Lulu Press, Inc, 2004. (ISBN 1-4116-0724-4.)

[10] Hickey S, Roberts H. Misleading information on the properties of vitamin C. *PLoS Med* 2005;2:e307.

[11] Day N, Oakes S, Luben R, Khaw KT, Bingham S, Welch A, et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer* 1999;80(suppl 1):95–103.

[12] Cox B, Huppert F, Whichelow M. The Health and Lifestyle Survey: seven years on. Aldershot, UK, Dartmouth Publishing Company, 1993.

[13] Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, et al. Validity and repeatability of a simple index derived from the short physical activity questionnaire used

in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Public Health Nutr* 2003;6:407–13.

[14] Elias P, Halstead K, Prandy K. CASOC: computer-assisted standard occupational coding. London, UK: HMSO, 1993.

[15] Shohaimi S, Luben R, Wareham N, Day N, Bingham S, Welch A, et al. Residential area deprivation predicts smoking habit independently of individual educational level and occupational social class. A cross sectional study in the Norfolk cohort of the European Prospective Investigation into Cancer (EPIC-Norfolk). *J Epidemiol Community Health* 2003;57:270–6.

[16] Townsend, P., Phillimore, P. and Beattie, A. (1988) Health and Deprivation: Inequality and the North. Routledge, London.

[17] Welch AA, Luben R, Khaw KT, Bingham SA. The CAFE computer program for nutritional analysis of the EPIC-Norfolk food frequency questionnaire and identification of extreme nutrient values. *J Hum Nutr Dietet* 2005;18:99–116.

[18] Vuilleumier J, Keck E. Fluorometric assay of vitamin C in biological materials using a centrifugal analyser with fluorescence attachment. J Micronutr Anal 1989;5:25–34.

[19] Sinha S, Myint PK, Luben RN, Khaw KT. Accuracy of death certification and hospital record linkage for identification of incident stroke. *BMC Med Res Methodol*. 2008;8:74.

[20] Roca J, Burgos F, Sunyer J, Saez M, Chinn S, Antó JM, et al. References values for forced spirometry. Group of the European Community Respiratory Health Survey. *Eur Respir* J. 1998;11:1354-62.

[21] Luo J, Shen L, Zheng D. Association between vitamin C intake and lung cancer: a dose-response meta-analysis. *Sci Rep.* 2014;4:6161.

[22] Wu TC, Huang YC, Hsu SY, Wang YC, Yeh SL. Vitamin E and vitamin C supplementation in patients with chronic obstructive pulmonary disease. *Int J Vitam Nutr Res*. 2007;77:272-9.

[23] Heart Protection Study Collaborative Group.MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet*. 2002;360:23-33.

[24] Hemilä H, Louhiala P. Vitamin C for preventing and treating pneumonia. *Cochrane Database Syst Rev*. 2013;8:CD005532.

[25] McKeever TM, Lewis SA, Smit HA, Burney P, Cassano PA, Britton J. A multivariate analysis of serum nutrient levels and lung function. *Respir Res.* 2008;9:67.

[26] Ochs-Balcom HM, Grant BJ, Muti P, Sempos CT, Freudenheim JL, Browne RW, et al. Antioxidants, oxidative stress, and pulmonary function in individuals diagnosed with asthma or COPD. *Eur J Clin Nutr.* 2006;60:991-9.

[27] Ford ES, Li C, Cunningham TJ, Croft JB. Associations between antioxidants and all-cause mortality among US adults with obstructive lung function. *Br J Nutr*. 2014;112:1662-73.

[28] Ahmadi A, Haghighat N, Hakimrabet M, Tolide-ie H. Nutritional evaluation in chronic obstructive pulmonary disease patients. *Pak J Biol Sci.* 2012;15:501-5.

[29] ben Anes A, Fetoui H, Bchir S, ben Nasr H, Chahdoura H, Chabchoub E, et al. Increased oxidative stress and altered levels of nitric oxide and peroxynitrite in Tunisian patients with chronic obstructive pulmonary disease: correlation with disease severity and airflow obstruction. *Biol Trace Elem Res*. 2014 ;161:20-31.

[30] Koike K, Ishigami A, Sato Y, Hirai T, Yuan Y, Kobayashi E, et al. Vitamin C prevents cigarette smoke-induced pulmonary emphysema in mice and provides pulmonary restoration. *Am J Respir Cell Mol Biol.* 2014;50:347-57.

[31] Li W, Maeda N, Beck MA. Vitamin C deficiency increases the lung pathology of influenza virus-infected gulo-/- mice. *J Nutr.* 2006;136:2611-6.

[32] Hamer DH, Sempértegui F, Estrella B, Tucker KL, Rodríguez A, Egas J, et al. Micronutrient deficiencies are associated with impaired immune response and higher burden of respiratory infections in elderly Ecuadorians. *J Nutr.* 2009;139:113-9.

441 **Table 1:** Baseline characteristics of men and women of EPIC-Norfolk 39-79 years old at the baseline by plasma vitamin C quartile categories

	Vitamin C quartiles				
Characteristics	Q1 (3-41 umol/l) N = 5068	Q2 (42-54 umol/l) N = 4849	Q3 (54.1-66 umol/l) N = 4868	Q4 (66-242 umol/l) N = 4554	p-for trend
Vit C in plasma (umol/l)	28.19 (9.69)	48.51 (3.71)	60.26 (3.38)	78.77 (13.01)	<0.001
Age (years)	60.07 (9.40)	58.93 (9.39)	58.28 (9.03)	58.53 (9.16)	<0.001
FVC (% pred)	87.99 (18.64)	91.43 (18.64)	94.57 (18.36)	96.32 (18.99)	<0.001
FEV1 (% pred)	88.90 (17.94)	92.42 (17.19)	94.81 (16.62)	95.78 (17.29)	<0.001
Townsend index	-1.82 (2.33)	-2.14 (2.07)	-2.19 (2.08)	-2.16 (2.08)	<0.001
Male	3294 (65.0%)	2629 (54.2%)	1924 (39.5%)	1190 (26.1%)	<0.001
Current smoker	1048 (20.8%)	464 (9.6%)	389 (8.1%)	344 (7.6%)	<0.001
Alcohol consumption (grams/week)	8.98 (14.56)	9.01 (13.28)	8.41 (12.02)	8.41 (11.75)	0.023
BMI > 30 kg/m^2	934 (18.5%)	814 (16.8%)	617 (12.7%)	403 (8.9%)	<0.001
Educational attainment					<0.001
None	2155 (42.5%)	1761 (36.4%)	1627 (33.4%)	1522 (33.4%)	
O-level	472 (9.3%)	498 (10.3%)	494 (10.2%)	525 (11.5%)	
A-Level	1988 (39.2%)	1974 (40.8%)	2016 (41.4%)	1813 (39.8%)	
Degree or higher	452 (8.9%)	611 (12.6%)	728 (15.0%)	693 (15.2%)	
Occupational social class					<0.001

Professional	270 (5.5%)	312 (6.6%)	365 (7.6%)	361 (8.1%)	
Managerial	1508 (30.6%)	1694 (35.7%)	1856 (38.8%)	1859 (41.6%)	
Skilled non-manual	723 (14.7%)	797 (16.8%)	837 (17.5%)	739 (16.5%)	
Skilled manual	1403 (28.5%)	1155 (24.4%)	993 (20.7%)	881 (19.7%)	
Semi-skilled	790 (16.0%)	636 (13.4%)	585 (12.2%)	527 (11.8%)	
Non-skilled	235 (4.8%)	149 (3.1%)	152 (3.2%)	102 (2.3%)	
Physically inactive	3139 (61.9%)	2859 (59.0%)	2749 (56.5%)	2501 (54.9%)	<0.001
Myocardial infarction (yes)	249 (4.9%)	151 (3.1%)	114 (2.3%)	81 (1.8%)	<0.001
Diabetes (yes)	169 (3.3%)	127 (2.6%)	78 (1.6%)	43 (0.9%)	<0.001
Stroke (yes)	100 (2.0%)	57 (1.2%)	52 (1.1%)	45 (1.0%)	<0.001
Outcomes					
Hospitalisations					
All (ex-lung cancer)	1237 (24.4%)	977 (20.1%)	865 (17.8%)	727 (16.0%)	<0.001
Lung cancer	162 (3.2%)	64 (1.3%)	49 (1.0%)	50 (1.1%)	<0.001
Chronic respiratory illnesses	503 (9.9%)	352 (7.3%)	307 (6.3%)	268 (5.9%)	<0.001

Pneumonia	463 (9.1%)	337 (6.9%)	285 (5.9%)	214 (4.7%)	<0.001
Mortality					
All (ex-lung cancer)	147 (2.9%)	104 (2.1%)	86 (1.8%)	70 (1.5%)	<0.001
Lung cancer	143 (2.8%)	56 (1.2%)	45 (0.9%)	36 (0.8%)	<0.001
Chronic respiratory illnesses	51 (1.0%)	23 (0.5%)	17 (0.3%)	14 (0.3%)	<0.001
Pneumonia	63 (1.2%)	48 (1.0%)	41 (0.8%)	29 (0.6%)	<0.001
Incidence (hospitalisation or mortality)					
All (ex-lung cancer)	1274 (25.1%)	1002 (20.7%)	889 (18.3%)	749 (16.5%)	<0.001
Lung cancer	190 (3.7%)	70 (1.4%)	54 (1.1%)	53 (1.2%)	<0.001
Chronic respiratory illnesses	517 (10.2%)	355 (7.3%)	311 (6.4%)	270 (5.9%)	<0.001
Pneumonia	489 (9.6%)	357 (7.4%)	302 (6.2%)	228 (5.0%)	<0.001

442 Data presented are mean (SD) for continuous and number (%) for categorical data. Q1 represents the lowest and the Q4 represents the highest quartiles
443 groups of plasma vitamin C concentrations.
444

445 **Table 2:** Hazard ratios and corresponding 95% confidence intervals for incident any respiratory illness (except lung cancers), lung cancer, chronic respiratory
446 illnesses and pneumonia by quartiles of plasma vitamin C concentration

	Model 1 (n=19,336):	Model 2 (n=18,890):	Model 3 (n=18,168):	Model 4 (n=17,744)	Model 5 (n=18,168)	Model 6 (n=17,744):	Model 7 (n=16,569)
Any respiratory Condition[1] (events =3914)							
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	0.81 (0.74,0.88)	0.84 (0.78,0.92)	0.93 (0.85,1.01)	0.94 (0.86,1.03)	0.93 (0.85,1.01)	0.95 (0.87,1.04)	0.95 (0.86,1.04)
Q3	0.73 (0.67,0.79)	0.77 (0.71,0.85)	0.87 (0.79,0.95)	0.88 (0.81,0.97)	0.87 (0.80,0.96)	0.89 (0.81,0.98)	0.89 (0.81,0.98)
Q4	0.67 (0.61,0.73)	0.73 (0.66,0.80)	0.82 (0.74,0.90)	0.84 (0.76,0.93)	0.83 (0.75,0.91)	0.85 (0.77,0.95)	0.86 (0.77,0.95)
Lung cancer[2] (events=367)							
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	0.40 (0.31,0.53)	0.42 (0.32,0.56)	0.58 (0.44,0.77)	0.59 (0.44,0.79)	0.58 (0.44,0.77)	0.59 (0.44,0.79)	0.61 (0.45,0.82)
Q3	0.34 (0.25,0.46)	0.38 (0.28,0.51)	0.51 (0.37,0.70)	0.53 (0.38,0.74)	0.51 (0.37,0.70)	0.53 (0.38,0.74)	0.55 (0.40,0.77)
Q4	0.38 (0.28,0.52)	0.43 (0.31,0.59)	0.55 (0.40,0.78)	0.58 (0.41,0.81)	0.55 (0.39,0.78)	0.57 (0.41,0.81)	0.53 (0.36,0.79)
Chronic Respiratory illnesses (events =1,453)							
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	0.68 (0.60,0.78)	0.75 (0.65,0.86)	0.88 (0.76,1.01)	0.90 (0.78,1.04)	0.88 (0.76,1.01)	0.90 (0.78,1.04)	0.90 (0.77,1.04)
Q3	0.58 (0.51,0.67)	0.67 (0.58,0.78)	0.82 (0.71,0.96)	0.85 (0.73,0.99)	0.83 (0.71,0.96)	0.85 (0.73,0.99)	0.85 (0.73,1.00)
Q4	0.54 (0.46,0.63)	0.64 (0.55,0.75)	0.78 (0.66,0.92)	0.81 (0.69,0.96)	0.78 (0.67,0.92)	0.81 (0.69,0.96)	0.80 (0.67,0.95)
Pneumonia (events =1,376)							
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	0.77 (0.67,0.88)	0.81 (0.70,0.93)	0.86 (0.74,0.99)	0.87 (0.75,1.01)	0.86 (0.74,0.99)	0.88 (0.76,1.02)	0.88 (0.75,1.02)
Q3	0.68 (0.59,0.79)	0.71 (0.61,0.82)	0.78 (0.67,0.91)	0.79 (0.68,0.93)	0.79 (0.68,0.92)	0.81 (0.69,0.94)	0.81 (0.69,0.95)
Q4	0.55 (0.47,0.65)	0.60 (0.51,0.71)	0.66 (0.56,0.79)	0.68 (0.57,0.81)	0.68 (0.57,0.81)	0.70 (0.59,0.83)	0.70 (0.58,0.85)

[1]includes J45-J46, J00-J99, J40-J47, J12-J18, B012, B052, B953, B960, B961, J100, J110, J851; [2]includes C33-C34. Q1 represents the lowest and the Q4 represents the highest quartiles groups of plasma vitamin C concentrations.
Model 1) age and sex; 2) age, sex, and respiratory function assessed using FVC and FEV1; 3) age, sex, respiratory function, and lifestyle factors (smoking, alcohol consumption, physical activity and BMI); 4) age, sex, respiratory function, lifestyle factors, occupational social class and Townsend index of deprivation; 5) as in model 3 with additional adjustment for prevalent diabetes, MI and stroke; 6) (fully adjusted model) as in model 4 with additional adjustment for history of diabetes, myocardial infarction and stroke; and finally model 7 |(sensitivity model) was constructed as in model 6 after excluding participants who were taking vitamin C containing supplements.

455 **Table 3:** Hazard ratios and corresponding 95%CI for all respiratory deaths, lung cancer deaths chronic respiratory cause as cause of death and pneumonia
456 deaths by quartiles of plasma vitamin C concentration

	Model 1 (n=19,336):	Model 2 (n=18,890):	Model 3 (n=18,168):	Model 4 (n=17,744)	Model 5 (n=18,168)	Model 6 (n=17,744):	Model 7 (n=16,569)
Any respiratory Condition[1] (events =407)							
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	0.77 (0.6,0.99)	0.78 (0.60,1.01)	0.89 (0.68,1.17)	0.93 (0.71,1.23)	0.89 (0.68,1.17)	0.93 (0.71,1.23)	0.95 (0.72,1.26)
Q3	0.69 (0.53,0.91)	0.74 (0.56,0.98)	0.89 (0.66,1.18)	0.93 (0.69,1.24)	0.89 (0.67,1.18)	0.93 (0.70,1.25)	0.96 (0.71,1.29)
Q4	0.59 (0.44,0.79)	0.68 (0.50,0.91)	0.79 (0.58,1.08)	0.81 (0.59,1.11)	0.80 (0.58,1.09)	0.81 (0.59,1.12)	0.87 (0.62,1.22)
Lung cancer[2] (events=280)							
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	0.44 (0.32,0.59)	0.46 (0.34,0.63)	0.63 (0.46,0.88)	0.64 (0.46,0.89)	0.63 (0.46,0.87)	0.64 (0.46,0.89)	0.65 (0.46,0.91)
Q3	0.39 (0.28,0.55)	0.44 (0.31,0.61)	0.58 (0.40,0.83)	0.60 (0.41,0.86)	0.58 (0.41,0.83)	0.60 (0.41,0.86)	0.61 (0.42,0.89)
Q4	0.37 (0.25,0.53)	0.40 (0.28,0.59)	0.52 (0.35,0.78)	0.54 (0.35,0.81)	0.52 (0.35,0.78)	0.54 (0.35,0.81)	0.48 (0.30,0.77)
Chronic Respiratory illnesses (events =105)							
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	0.50 (0.30,0.81)	0.56 (0.34,0.93)	0.87 (0.52,1.47)	0.96 (0.56,1.64)	0.87 (0.51,1.46)	0.95 (0.56,1.63)	0.93 (0.54,1.6)
Q3	0.40 (0.23,0.70)	0.47 (0.27,0.84)	0.73 (0.40,1.35)	0.83 (0.45,1.54)	0.74 (0.40,1.35)	0.84 (0.45,1.56)	0.81 (0.43,1.53)
Q4	0.38 (0.20,0.69)	0.55 (0.30,1.02)	0.73 (0.38,1.40)	0.84 (0.43,1.64)	0.74 (0.38,1.41)	0.85 (0.44,1.66)	0.63 (0.29,1.41)
Pneumonia (events =181)							
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	0.82 (0.56,1.19)	0.79 (0.53,1.16)	0.78 (0.52,1.18)	0.85 (0.56,1.28)	0.78 (0.52,1.18)	0.85 (0.56,1.29)	0.86 (0.56,1.32)
Q3	0.76 (0.51,1.14)	0.79 (0.53,1.18)	0.84 (0.55,1.29)	0.90 (0.59,1.38)	0.85 (0.55,1.29)	0.91 (0.60,1.40)	0.97 (0.63,1.50)

Q4	0.54 (0.35,0.86)	0.59 (0.37,0.94)	0.64 (0.40,1.02)	0.61 (0.37,1.00)	0.64 (0.40,1.04)	0.61 (0.37,1.01)	0.66 (0.39,1.10)

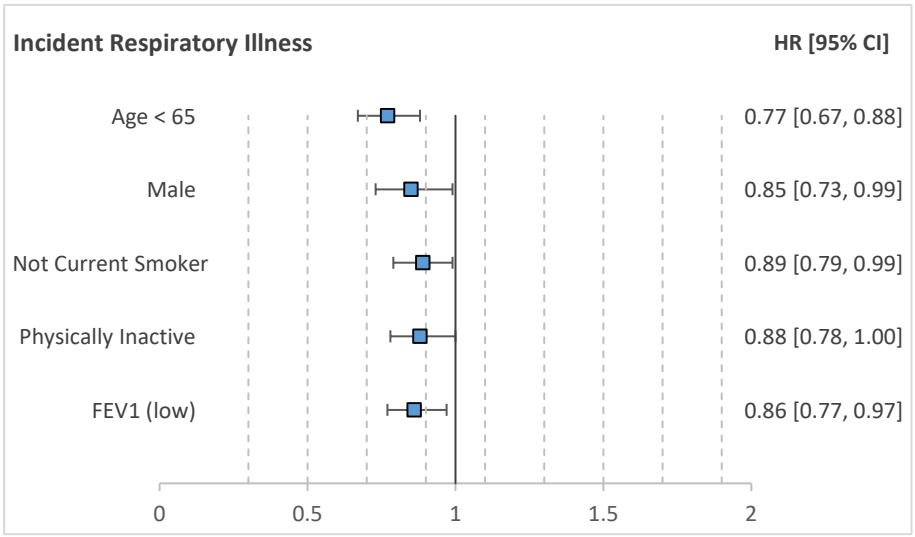
[1]includes J45-J46, J00-J99, J40-J47, J12-J18, B012, B052, B953, B960, B961, J100, J110, J851; [2]includes C33-C34. Q1 represents the lowest and the Q4 represents the highest quartiles groups of plasma vitamin C concentrations.

Model 1) age and sex; 2) age, sex, and respiratory function assessed using FVC and FEV1; 3) age, sex, respiratory function, and lifestyle factors (smoking, alcohol consumption, physical activity and BMI); 4) age, sex, respiratory function, lifestyle factors, occupational social class and Townsend index of deprivation; 5) as in model 3 with additional adjustment for prevalent diabetes, MI and stroke; 6) (fully adjusted model) as in model 4 with additional adjustment for history of diabetes, myocardial infarction and stroke; and finally model 7 (sensitivity model) was constructed as in model 6 after excluding participants who were taking vitamin C containing supplements.

469

Figure 1: Hazard ratios for study outcomes in people in the highest quartile group (Q4) compared to the lowest quartile group (Q1) stratified by important factors for model 6.

Figure 1 (a): Incident respiratory illness



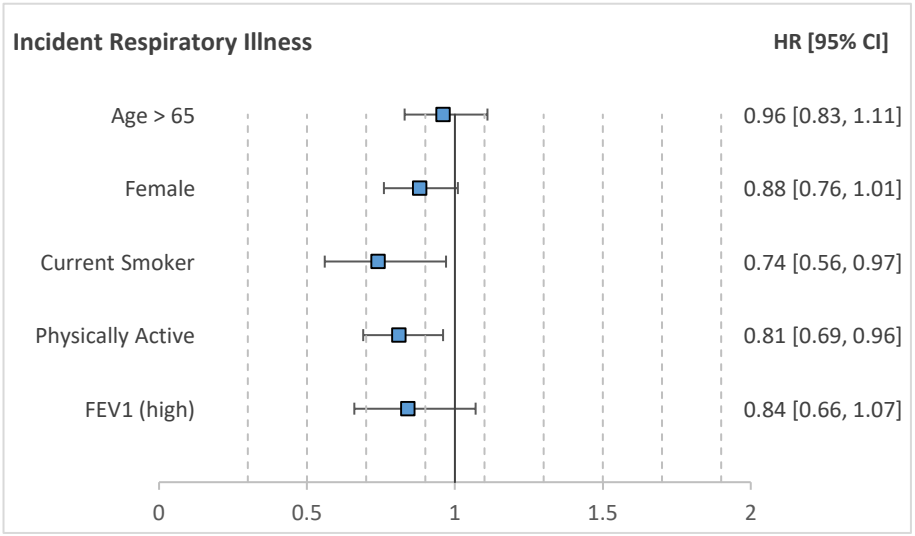
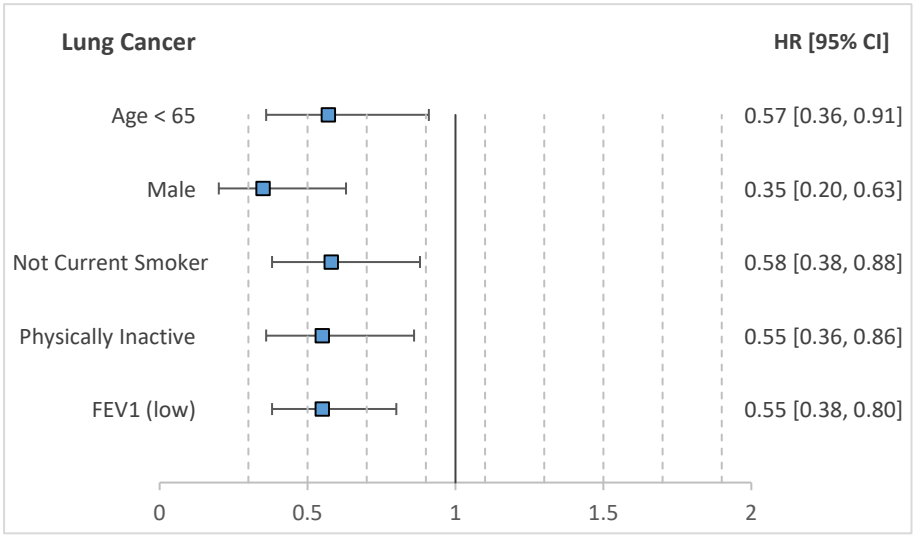


Figure 1 (b): Incident lung cancer



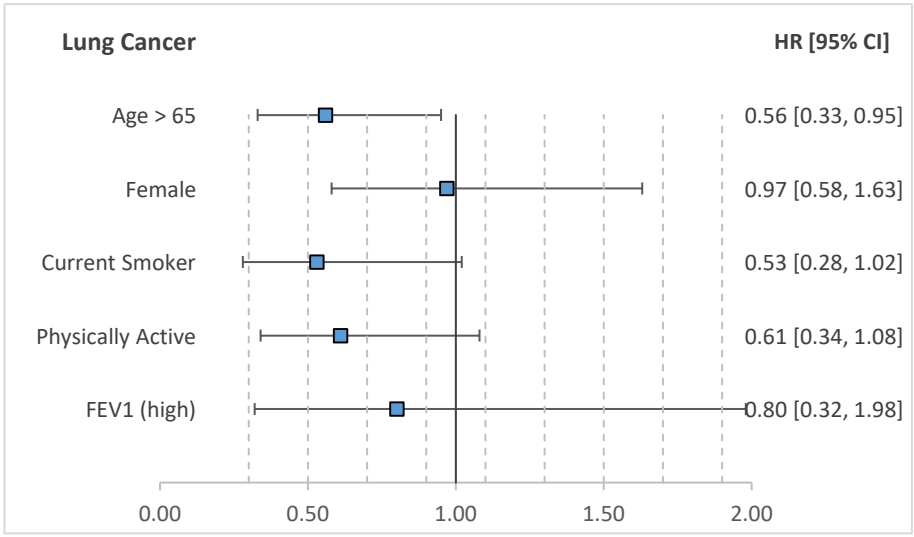
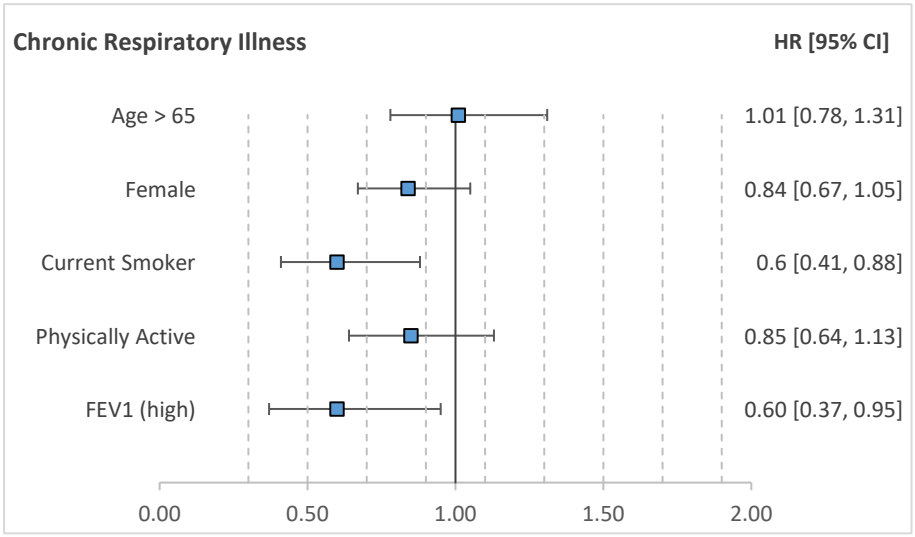


Figure 1 (c): Incident chronic respiratory diseases



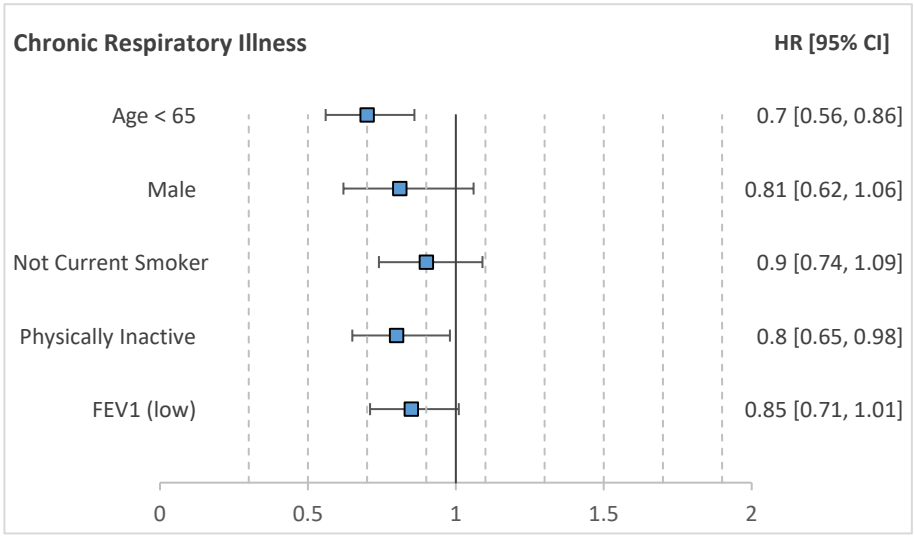
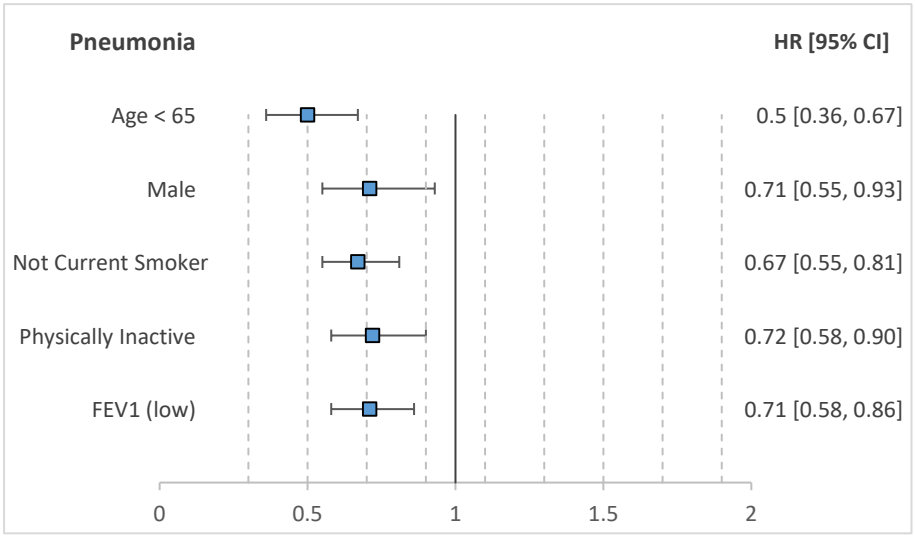
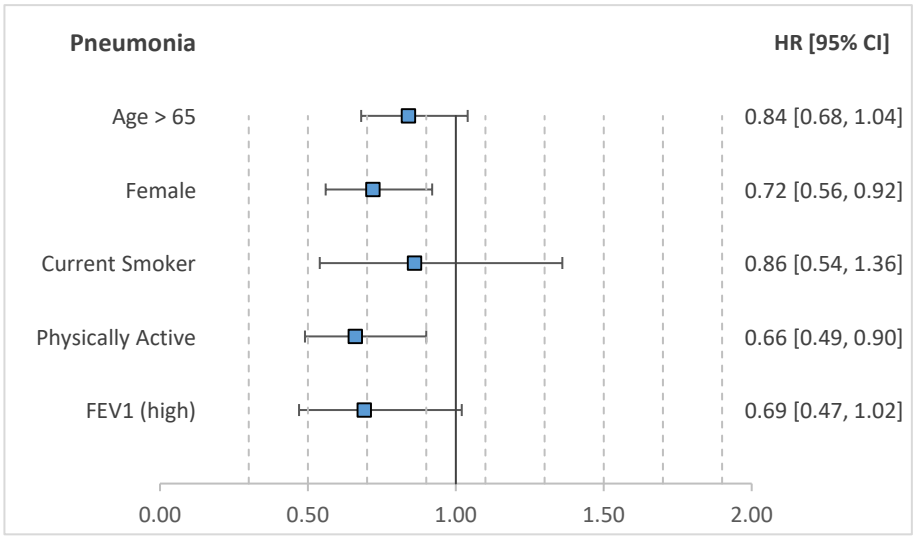
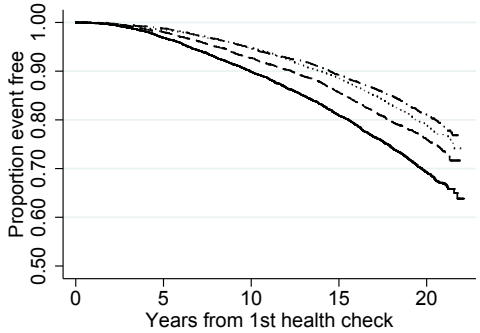
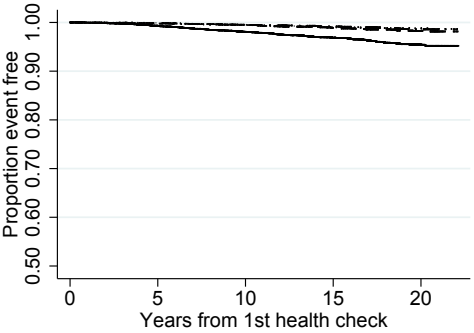


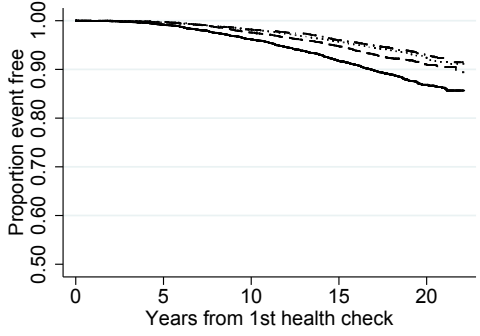
Figure 1 (d): Incident pneumonia

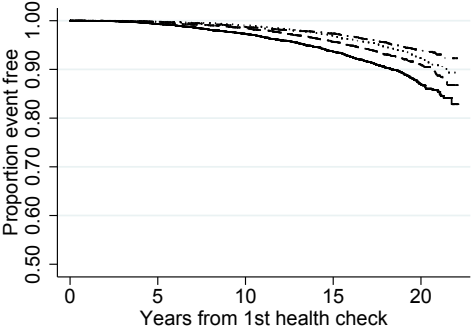





Proportion event free
0.50 0.60 0.70 0.80 0.90 1.00
0 5 10 15 20
Years from 1st health check


Proportion event free
0.50 0.60 0.70 0.80 0.90 1.00
0 5 10 15 20
Years from 1st health check


Proportion event free
0.50 0.60 0.70 0.80 0.90 1.00
0 5 10 15 20
Years from 1st health check


Proportion event free
0.50 0.60 0.70 0.80 0.90 1.00
0 5 10 15 20
Years from 1st health check

Supplementary Material: Measurement methods

The participants were asked about their medical history with the question, “Has a doctor ever told you that you have any of the following?”, in reference to the following conditions: stroke, heart attack, diabetes, cancer and respiratory diseases (asthma/COPD). Smoking history was obtained by asking the following yes or no questions: “Have you ever smoked as much as one cigarette a day for as long as a year?” and “Do you smoke cigarettes now?”. Participants who took any supplements or supplements containing vitamin C were identified from the question, “Have you taken any vitamins, minerals, or other food supplements regularly during the past year (such as vitamin C, vitamin D, iron, calcium, fish oils, primrose oil, β-carotene)?”, which was followed by “name/brand and dose.”

Height and weight were measured, and body mass index (BMI) was calculated as weight (kg) divided by height squared (m2). Blood pressure (BP) was measured with an Accutorr monitor (Datascope, Huntingdon, United Kingdom) after the participant had been seated for 5 min. We used the mean of 2 measurements for analysis. Respiratory function was assessed by spirometry [12]. Spirometry (for measurement of FEV1 and FVC) was performed twice to ensure there were two acceptable flow volume loops using a portable spirometer (Micro medical, Rochester, UK), the higher of the two measures from either of the forced exhalations was used for analyses.

A 4-level physical activity index was derived from the validated EPIC short physical activity questionnaire designed to assess combined work and leisure activity. Participants were categorized into inactive, moderately inactive, moderately active, and active categories. The validity and repeatability of this scoring system was detailed elsewhere [13]. We combined these into inactive (inactive and moderately inactive) and active (moderately active and

active) categories. Social class was classified according to the Registrar General's occupation-based classification scheme [14]. Social class I consists of professionals, social class II includes managerial and technical occupations, social class III is subdivided into nonmanual skilled workers and manual skilled workers, social class IV consists of partly skilled workers, and social class V comprises unskilled manual workers [15]. Social class was also combined as nonmanual (I, II, and III nonmanual) and manual (III manual, IV, and V) social classes. In addition we also used the Townsend deprivation index (TDI) as a measure of material deprivation; it takes into account of unemployment, car and home ownership within a household [16].

Educational status was recorded as no qualification, O- level, A-level, degree or higher qualification. Educational attainment was re-categorised as low educational attainment (no or O level) and high educational attainment (at least A level). Alcohol consumption was derived from a food-frequency questionnaire (FFQ) collected at baseline. For the “drinks” category, responses ranging from never to >6 times/d were given for 4 types of alcoholic drinks. An in-house computer program, CAFE, was developed for data entry and analysis [17].

Supplementary Table 1: Hazard ratios (95%CI) for mortality respiratory illness, lung cancer, chronic respiratory diseases and pneumonia stratified by important factors for model 6.

	Age <65	Age 65 or more	Male	Female	Not Current smoker	current smoker	Physically Inactive	Physically Active	FEV (low)	FEV (high)[3]
Any respiratory Condition[1] (events =407)										
N (events)	91	273	200	164	290	74	266	98	310	54
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	1.03 (0.61,1.73)	0.89 (0.64,1.23)	0.96 (0.67,1.37)	0.92 (0.59,1.44)	0.93 (0.69,1.27)	0.95 (0.51,1.77)	0.87 (0.63,1.21)	1.16 (0.67,2.00)	0.96 (0.71,1.29)	0.97 (0.45,2.07)
Q3	0.82 (0.45,1.49)	0.99 (0.71,1.38)	1.03 (0.7,1.52)	0.88 (0.56,1.38)	0.95 (0.69,1.31)	0.96 (0.46,2.01)	0.84 (0.6,1.19)	1.27 (0.72,2.24)	1 (0.73,1.37)	0.7 (0.31,1.55)
Q4	0.53 (0.25,1.12)	0.91 (0.64,1.30)	0.76 (0.46,1.26)	0.86 (0.55,1.34)	0.83 (0.58,1.18)	0.82 (0.35,1.89)	0.81 (0.56,1.17)	0.85 (0.44,1.64)	0.81 (0.57,1.16)	0.79 (0.38,1.68)
Lung cancer[2] (events=280)										
N (events)	132	125	180	77	154	103	164	93	230	27
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	0.74 (0.48,1.15)	0.53 (0.32,0.88)	0.58 (0.4,0.86)	0.91 (0.47,1.76)	0.54 (0.35,0.82)	0.85 (0.5,1.43)	0.65 (0.43,0.98)	0.6 (0.35,1.05)	0.59 (0.42,0.84)	1.14 (0.42,3.11)
Q3	0.56 (0.33,0.93)	0.64 (0.38,1.08)	0.47 (0.29,0.77)	1.03 (0.55,1.95)	0.61 (0.4,0.94)	0.48 (0.23,1.01)	0.63 (0.4,0.99)	0.55 (0.29,1.02)	0.56 (0.38,0.83)	0.96 (0.33,2.77)
Q4	0.49 (0.27,0.89)	0.57 (0.32,1.01)	0.38 (0.2,0.71)	0.92 (0.49,1.75)	0.47 (0.28,0.78)	0.69 (0.35,1.39)	0.53 (0.31,0.89)	0.54 (0.28,1.06)	0.55 (0.36,0.85)	0.52 (0.15,1.82)
Chronic Respiratory condition (events =105)										

N (events)	35	62	63	34	50	47	68	29	92	5
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	1.24 (0.55,2.81)	0.76 (0.38,1.53)	1.15 (0.60,2.20)	0.58 (0.22,1.53)	0.78 (0.38,1.62)	1.14 (0.54,2.42)	0.79 (0.41,1.52)	1.69 (0.65,4.41)	-	-
Q3	1.27 (0.49,3.28)	0.62 (0.27,1.42)	0.92 (0.41,2.07)	0.68 (0.26,1.79)	0.8 (0.36,1.78)	0.86 (0.32,2.29)	0.69 (0.33,1.47)	1.53 (0.50,4.70)	-	-
Q4	0.36 (0.08,1.71)	1.14 (0.54,2.41)	1.17 (0.50,2.74)	0.59 (0.21,1.65)	0.92 (0.40,2.12)	0.79 (0.26,2.38)	0.7 (0.32,1.55)	1.39 (0.42,4.61)	-	-
Pneumonia (events=158)										
N (events)	30	128	81	77	141	17	121	37	131	31
Q1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	1.28 (0.53,3.06)	0.77 (0.48,1.24)	0.91 (0.53,1.58)	0.76 (0.39,1.45)	0.88 (0.57,1.36)	0.56 (0.12,2.66)	0.72 (0.45,1.17)	1.47 (0.60,3.64)	0.81 (0.52,1.28)	0.96 (0.36,2.52)
Q3	0.96 (0.36,2.60)	0.93 (0.58,1.49)	1.05 (0.59,1.88)	0.8 (0.42,1.51)	0.93 (0.59,1.46)	0.85 (0.18,4.08)	0.87 (0.54,1.42)	1.27 (0.48,3.37)	1.03 (0.65,1.61)	0.52 (0.17,1.58)
Q4	0.16 (0.02,1.29)	0.71 (0.42,1.20)	0.5 (0.21,1.21)	0.64 (0.33,1.21)	0.58 (0.34,0.98)	1.54 (0.38,6.20)	0.62 (0.36,1.09)	0.67 (0.22,2.03)	0.6 (0.34,1.05)	0.63 (0.24,1.67)

[1]includes J45-J46, J00-J99, J40-J47, J12-J18, B012, B052, B953, B960, B961, J100, J110, J851; [2]includes C33-C34.

[3]High FEV is defined as the upper quartile and low as the remaining. Q1 represents the lowest and the Q4 represents the highest quintiles groups of plasma vitamin C concentrations.

Note due to the number of variables included in the models, some stratified analyses were not reported because number of events are less than required minimum number of events or two low.

Supplementary Table 2: Hazard ratios (95%CI) for incidence and mortality of respiratory illness, lung cancer, chronic respiratory diseases and pneumonia in non-smokers and never-smokers

		Incidence					**Mortality**		
	Model 6 (non-smokers)	Model 6 (never-smokers)	Model 7 (non-smokers)	Model 7 (never-smokers)		Model 6 (non-smokers)	Model 6 (never-smokers)	Model 7 (non-smokers)	Model 7 (never-smokers)
Any respiratory Condition									
Q1	1.00	1.00	1.00	1.00		1.00	1.00	1.00	1.00
Q2	0.98 (0.89,1.08)	1.04 (0.89,1.22)	0.98 (0.89,1.08)	1.04 (0.88,1.22)		0.93 (0.69,1.27)	0.98 (0.56,1.71)	0.97 (0.7,1.32)	1.09 (0.61,1.95)
Q3	0.91 (0.82,1.01)	0.92 (0.78,1.08)	0.91 (0.82,1.01)	0.92 (0.78,1.08)		0.95 (0.69,1.31)	1.24 (0.73,2.12)	0.98 (0.7,1.36)	1.33 (0.75,2.36)
Q4	0.89 (0.79,0.99)	0.82 (0.69,0.97)	0.89 (0.79,1)	0.81 (0.67,0.96)		0.83 (0.58,1.18)	0.97 (0.55,1.71)	0.91 (0.63,1.31)	1.14 (0.62,2.10)
Lung cancer									
Q1	1.00	1.00	1.00	1.00		1.00	1.00[1]	1.00	1.00[1]
Q2	0.49 (0.34,0.72)	0.58 (0.18,1.83)	0.51 (0.35,0.75)	0.60 (0.19,1.9)		0.54 (0.35,0.82)	1.03 (0.27,3.88)	0.56 (0.37,0.86)	1.09 (0.29,4.11)
Q3	0.53 (0.36,0.79)	0.93 (0.34,2.54)	0.55 (0.37,0.83)	0.75 (0.26,2.18)		0.61 (0.4,0.94)	1.31 (0.37,4.59)	0.64 (0.41,0.99)	0.99 (0.26,3.83)
Q4	0.58 (0.38,0.88)	1.11 (0.41,3.01)	0.60 (0.39,0.94)	0.91 (0.31,2.66)		0.47 (0.28,0.78)	0.86 (0.22,3.39)	0.47 (0.27,0.83)	0.39 (0.07,2.20)
Chronic Respiratory illnesses									
Q1	1.00	1.00	1.00	1.00		1.00	1.00[2]	1.00	1.00[2]
Q2	0.94 (0.79,1.12)	0.99 (0.74,1.34)	0.92 (0.78,1.10)	0.96 (0.71,1.30)		0.76 (0.37,1.59)	0.95 (0.05,16.89)	0.80 (0.38,1.68)	1.08 (0.06,18.77)
Q3	0.91 (0.76,1.09)	1.01 (0.75,1.37)	0.91 (0.76,1.09)	1.02 (0.76,1.38)		0.82 (0.37,1.81)	1.45 (0.09,24.45)	0.79 (0.35,1.80)	1.73 (0.10,29.65)

Q4	0.90 (0.74,1.09)	0.95 (0.70,1.29)	0.88 (0.72,1.07)	0.88 (0.64,1.21)		0.94 (0.41,2.18)	4.41 (0.45,43.20)	0.65 (0.23,1.83)	2.14 (0.16,28.18)
Pneumonia									
Q1	1.00	1.00	1.00	1.00		1.00	1.00[2]	1.00	1.00[2]
Q2	0.85 (0.72,1.00)	0.84 (0.65,1.09)	0.84 (0.71,0.99)	0.83 (0.64,1.08)		0.88 (0.57,1.36)	0.71 (0.33,1.52)	0.87 (0.56,1.36)	0.7 (0.32,1.53)
Q3	0.76 (0.64,0.90)	0.75 (0.58,0.98)	0.75 (0.63,0.89)	0.77 (0.59,1.01)		0.93 (0.59,1.46)	1.26 (0.64,2.47)	0.98 (0.62,1.54)	1.32 (0.65,2.66)
Q4	0.67 (0.55,0.81)	0.59 (0.44,0.78)	0.68 (0.55,0.82)	0.59 (0.43,0.80)		0.58 (0.34,0.98)	0.56 (0.25,1.28)	0.63 (0.36,1.09)	0.62 (0.26,1.47)

[1]Excluded DM and CVA as could not be estimated from model due to lack of events.

[2]Excludes adjustment for BMI, activity, Deprivation, DM, MI and CVA due to lack of events.

Q1 represents the lowest and the Q4 represents the highest quintiles groups of plasma vitamin C concentrations.