Viewpoint

Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence?

Debbie A Lawlor, George Davey Smith, K Richard Bruckdorfer, Devi Kundu, Shah Ebrahim

In observational studies, antioxidant vitamins have been inversely associated with cardiovascular disease, cancer, and all-cause mortality.[1–3] However, well conducted randomised controlled trials have shown that supplementation with antioxidants does not protect against these disorders.[4] The figure, for example, summarises the effects of vitamin C seen in the largest randomised controlled trial to date,[5] and in a large observational study.[1] Both these papers were published in *The Lancet*, both are thought to be methodologically sound, and both are widely cited, yet their conclusions are contradictory. In this example, in the randomised trial vitamin C was part of a multivitamin supplement, whereas in the observational study plasma concentrations of vitamin C were assessed. However, it is difficult to see why a combination of vitamin C with other vitamins should reduce its protective effects, if they were real. Furthermore, in the case of other antioxidant vitamins, notably vitamin E, results of single-factor trials and observational studies show a similar discordance to those seen in our figure for vitamin C.[2–4]

Why did observational studies and randomised controlled trials come up with different answers?

Several reasons have been proposed for the disparity between the results of observational epidemiological studies and trials. First, antioxidants might be useful only for primary prevention of cardiovascular disease, and not protective once atherosclerosis is established.[6] However, trials[4,7] found that antioxidants did not reduce cardiovascular disease risk in participants who had no evidence of this disorder at randomisation. Second, in many of the trials the choice of antioxidant regimen has been criticised.[6] However, several trials have used doses that are broadly equivalent to levels seen in the observational studies: the comparisons in our figure are for similar concentrations of vitamin C in plasma. Further, the protective effects associated with high-concentration vitamin E supplements (about 200–400 IU per day) seen in the Nurses' Health and Physicians' observational studies (both with relative risks for cardiovascular mortality of less than 0·7)[2,3] were not seen in trials with lower (50 IU per day)[4] or similar (300 IU per day)[7] supplementation regimens. Third, the duration of most trials has been suggested to be inadequate to show benefit. However, in the Physicians' Health randomised trial,[8] no benefits were seen for cancer or cardiovascular disease outcomes during a mean follow-up period of 12 years.

A more likely explanation is that associations between antioxidants and disease in observational studies are confounded by social and behavioural factors acting across the life course.[9] Typically, statistical adjustments are made for single measures of potential confounders at one time point in the life course. This approach probably inadequately captures the full extent of the complex ways in which social and behavioural factors confound associations between vitamins and disease.

To explore this proposition in more detail, we have assessed the association of a wide range of indicators of socioeconomic position across the life course, anthropometric indicators of childhood environmental circumstances,[10]—and behavioural risk factors with plasma vitamin C and E status in the British Women's Heart and Health Study. This study consists of 4286 women aged 60–79 years who were randomly selected from 23 British towns. The study methods have been described in detail previously.[11]

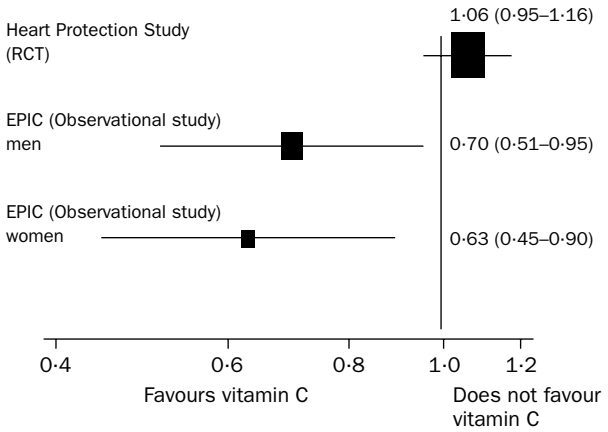
Tables 1 and 2 show the associations between life course socioeconomic position indicators, adult behavioural factors, and biomarkers of childhood environment with plasma vitamin C and E concentrations. All ten indicators of socioeconomic position from across the life course were linearly associated with vitamin C and most with vitamin E, such that people from poorer socioeconomic positions at any time had lower vitamin concentrations. With mutual adjustment for each of the other socioeconomic position indicators, most maintained their independent associations. Socioeconomic disadvantage had a cumulative effect over the life course, shown by a strong linear trend across a composite score of the ten socioeconomic indicators: in analysis adjusted for age and month of blood sampling, the odds of being in the highest quarter of the plasma vitamin C distribution decreased by 22% for each additional marker of adverse life course socio-economic position (odds ratio 0·88 [95% CI 0·84–0·92], p for linear trend <0·0001). Results for vitamin E were similar (0·92 [0·88–0·96], p for linear trend <0·0001). The associations with each vitamin where independent of those with the other vitamin.

Adult behavioural factors were also associated with vitamin status (tables 1 and 2). Women who smoked and those who were obese had lower vitamin C and E concentrations. Those who participated in leisure time exercise for at least 1 h per week, reported eating a low-fat diet or a high-fibre diet, and those who consumed alcohol daily had higher vitamin concentrations. These associations were independent of socioeconomic position across the life course. Markers of childhood development were associated with vitamin status in adult life. Women with longer legs and leg relative to trunk length, had

Lancet 2004; **363:** 1724–27
See Commentary page 1660

Department of Social Medicine, University of Bristol, UK (D A Lawlor PhD, Prof G Davey Smith DSc, Prof S Ebrahim DM); **and Department of Biochemistry and Molecular Biochemistry, Royal Free Hospital School of Medicine, University of London, UK** (Prof K R Bruckdorfer DSc, D Kundu PhD)

Correspondence to: Dr Debbie A Lawlor, Department of Social Medicine, University of Bristol, Canynge Hall, Whiteladies Road, Bristol, BS8 2PR
(e-mail: d.a.lawlor@bristol.ac.uk)



Estimates of the effects of an increase of 15·7 μmol/L plasma vitamin C on coronary heart disease 5-year mortality

Estimates from men and women from a large observational epidemiological (EPIC) study[3] and from a large randomised controlled trial (Heart Protection Study).[7] In the EPIC study, results were adjusted for age, systolic blood pressure, cholesterol, body-mass index, smoking, diabetes, and vitamin supplement use.

higher concentrations of vitamin C and E, independent of life course socioeconomic position and behavioural risk factors.

The conflicting observational and trial findings are probably the result of residual confounding caused by inadequate adjustment for the complexity of social and environmental exposures acting across the life course. We have shown here that vitamin status in adulthood is strongly associated with measures of socioeconomic position from childhood right through to adulthood, with anthropometric indicators of childhood environmental exposures and with adult behavioural risk factors. In a prospective cohort study death from cardiovascular disease was affected, in a cumulative fashion, by socioeconomic position and behavioural factors acting throughout the life course.[12] In that study the risk of death from cardiovascular disease was four times greater in those with the most adverse status for all socioeconomic and behavioural life course exposures compared with those with the most advantageous status for all exposures.

Better design of observational studies

Randomised controlled trials provide the most robust estimates of causal effects. However, a more careful approach to the implementation and interpretation of observational studies is needed, since trials are not always feasible and observational studies often generate hypotheses that are later tested in randomised controlled trials. In view of the expense of randomised controlled trials, only candidate agents with a high probability of being causal factors can be tested in this way. Here we outline ways in which the design and analysis of observational studies could be improved.

Appropriate adjustment for available confounding factors

In one of the largest prospective observational studies to find a beneficial effect of vitamin C on all-cause mortality, cardiovascular disease, and cancer, no adjustment was made for social class and physical activity. Although these data were available, they were not coded or analysed, because of resource limitations.[1] It would be instructive to reanalyse these data with full adjustment for these variables. However, results based on adjustment for a single measure of adulthood socioeconomic position should still be interpreted cautiously. Sensitivity analyses, as outlined below, would be informative in this situation.[13]

Measurement error and study design

The importance of measuring exposure and confounder variables accurately and precisely is underappreciated.[14] In

	Vitamin C quartile (range μmol/L)				Odds ratio (95% CI)*	p
	1 (0·00–20·46)	2 (20·47–39·27)	3 (39·28–59·64)	4 (59·65–190·47)		
Socioeconomic indicators						
Childhood						
Manual social class (%)	91·6 (89·3–93·4)	88·5 (86·0–90·7)	83·9 (81·0–86·5)	82·1 (79·2–84·7)	0·88 (0·84–0·92)	<0·0001
No bathroom in house (%)	45·1 (41·7–48·4)	40·7 (37·5–44·0)	33·7 (30·6–36·9)	33·4 (30·3–36·6)	0·92 (0·90–0·95)	<0·0001
No hot water in house (%)	39·5 (36·2–42·8)	37·5 (34·3–40·8)	31·7 (28·6–34·8)	29·9 (26·9–33·0)	0·93 (0·91–0·97)	<0·0001
Shared bedroom (%)	57·6 (54·2–60·9)	56·0 (52·6–59·3)	49·9 (46·6–53·2)	47·9 (44·6–51·3)	0·93 (0·90–0·97)	<0·0001
No car access (%)	86·5 (84·0–88·7)	84·7 (82·1–87·0)	79·9 (77·1–82·5)	78·9 (76·0–81·6)	0·91 (0·87–0·94)	<0·0001
Completed full-time education by age 18 years (%)	94·7 (93·0–96·0)	90·2 (88·0–92·1)	84·1 (81·5–86·4)	83·4 (80·8–85·7)	0·82 (0·78–0·86)	<0·0001
Adult						
Manual social class (%)	61·7 (58·2–65·0)	56·1 (52·6–59·4)	46·0 (42·6–49·3)	44·7 (41·4–48·0)	0·88 (0·85–0·91)	<0·0001
Local authority housing (%)	19·1 (16·7–21·8)	11·2 (9·3–13·4)	10·3 (8·5–12·5)	8·5 (6·9–10·5)	0·85 (0·81–0·88)	<0·0001
No car access (%)	35·8 (32·5–39·2)	27·3 (24·4–30·5)	22·6 (19·9–25·5)	22·6 (19·9–25·5)	0·90 (0·87–0·93)	<0·0001
State pension only (%)	36·7 (33·5–40·1)	28·7 (25·7–31·9)	22·3 (19·6–25·3)	24·4 (21·6–27·4)	0·89 (0·86–0·92)	<0·0001
Behavioural and lifestyle risk factors						
Current smoker (%)	17·6 (15·3–20·2)	11·5 (9·6–13·7)	7·5 (6·0–9·4)	6·4 (5·0–8·1)	0·81 (0·78–0·85)	<0·0001
Ex smoker (%)†	38·5 (35·1–42·0)	38·5 (35·2–41·9)	37·0 (33·9–40·4)	34·9 (31·8–38·2)	0·97 (0·94–1·00)	0·05
⩾1 h leisure exercise per week (%)	11·1 (9·3–13·3)	18·0 (15·6–20·7)	20·5 (18·0–23·3)	22·9 (20·2–25·8)	1·14 (1·10–1·19)	<0·0001
Low-fat diet (%)	13·5 (11·4–15·8)	13·9 (11·8–16·3)	15·9 (13·6–18·4)	19·7 (17·2–22·4)	1·08 (1·04–1·12)	<0·0001
High-fibre diet (%)	2·6 (1·7–3·8)	3·9 (2·8–5·3)	5·5 (4·2–7·1)	4·8 (3·6–6·4)	1·10 (1·02–1·18)	0·01
Obese: BMI >30 kg/m² (%)	31·5 (28·6–34·6)	28·0 (25·3–31·1)	24·6 (21·9–27·5)	21·1 (18·6–23·9)	0·92 (0·89–0·95)	<0·0001
Daily alcohol consumption (%)	10·9 (9·0–13·1)	16·1 (13·8–18·6)	18·7 (16·3–21·3)	22·0 (19·5–24·9)	1·15 (1·10–1·19)	<0·0001
Biomarkers childhood environment						
Adult height (mm)	1581·0 (1577·1–1584·9)	1583·3 (1581·4–1589·2)	1591·4 (1587·5–1595·3)	1594·3 (1590·4–1598·3)	2·20 (1·37–3·04)‡	<0·0001
Adult leg length (mm)	753·1 (750·4–755·8)	755·4 (752·7–758·0)	758·3 (752·7–758·0)	762·6 (759·9–765·3)	1·47 (0·90–2·05)‡	<0·0001
Adult leg to trunk ratio %	91·1 (90·7–91·4)	91·1 (90·8–91·5)	91·2 (90·9–91·6)	91·8 (91·5–92·2)	0·10 (0·02–0·17)‡	0·01

Data are prevalence (95% CI) for socioeconomic indicators and behavioural and lifestyle risk factors, and means (95% CI) for biomarkers childhood environment. All estimates adjusted for age and month of blood sampling (so seasonal effects are controlled for). *Per increase in 1 SD vitamin C. †Ex smoker=compared with never smokers, does not include current smokers. ‡Difference per increase in 1 SD vitamin C.

Table 1: **Life course factors across quartiles of plasma vitamin C**

	Vitamin E quartile (range μmol/L)				Odds ratio (95% CI)*	p
	1 (0·00–36.12)	2 (36·13–44·24)	3 (44·25–53·83)	4 (53·84–150·37)		
Socioeconomic indicators						
Childhood						
Manual social class (%)	88·6 (85·9–90·8)	90·3 (87·8–92·3)	85·9 (83·1–88·4)	82·1 (79·0–84·8)	0·82 (0·74–0·91)	<0·0001
No bathroom in house (%)	39·2 (35·9–42·7)	38·7 (35·4–42·2)	38·3 (35·0–41·7)	36·2 (33·0–39·7)	0·94 (0·87–0·99)	0·05
No hot water in house (%)	35·2 (31·9–38·6)	35·5 (32·2–38·9)	34·8 (31·5–38·2)	31·5 (28·4–34·9)	0·93 (0·86–0·99)	0·04
Shared bedroom (%)	55·4 (51·8–58·8)	54·5 (51·0–58·0)	55·0 (51·6–58·4)	49·4 (45·9–52·9)	0·89 (0·84–0·96)	0·002
No car access (%)	83·0 (80·1–85·5)	84·1 (81·4–86·6)	82·8 (80·0–85·3)	79·3 (76·3–82·0)	0·87 (0·80–0·95)	0·002
Completed full-time education by age 18 years (%)	90·8 (88·5–92·6)	89·5 (87·1–91·5)	87·8 (85·3–90·0)	84·5 (81·8–86·9)	0·76 (0·69–0·83)	<0·0001
Adult						
Manual social class (%)	55·3 (51·7–58·8)	55·2 (51·7–58·7)	53·6 (50·0–57·1)	46·3 (42·7–49·8)	0·83 (0·78–0·90)	<0·0001
Local authority housing (%)	13·1 (10·9–15·5)	12·3 (10·3–14·7)	15·2 (12·9–17·8)	10·3 (8·4 ,12·6)	0·91 (0·82–1·00)	0·08
No car access (%)	27·4 (24·4–30·7)	26·7 (23·6–29·9)	27·0 (23·9–30·2)	26·6 (23·5–29·8)	0·96 (0·88–1·04)	0·29
State pension only (%)	27·9 (24·8–31·3)	31·5 (28·2–34·9)	28·8 (25·7–32·1)	26·2 (23·1–29·4)	0·94 (0·87–1·00)	0·06
Behavioural risk factors						
Adult						
Current smoker (%)	13·0 (10·9–15·5)	10·8 (8·9–13·1)	10·0 (8·2–12·3)	9·7 (7·9–11·9)	0·88 (0·78–0·98)	0·02
Ex smoker (%)†	37·5 (34·0–41·1)	40·0 (36·4–43·5)	37·3 (33·9–40·8)	34·9 (31·6–38·4)	0·94 (0·88–1·02)	0·14
>1 h leisure exercise per week (%)	16·2 (13·8–18·8)	18·0 (15·6–20·8)	20·3 (17·7–23·2)	19·1 (16·5–21·9)	1·10 (1·02–1·20)	0·01
Low-fat diet (%)	13·9 (11·7–16·4)	15·0 (12·8–17·6)	15·6 (13·3–18·3)	16·1 (13·8–18·8)	1·05 (0·96–1·15)	0·33
High-fibre diet (%)	4·0 (2·9–5·6)	3·5 (2·5–5·0)	4·3 (3·2–6·0)	5·3 (4·0–7·0)	1·13 (0·98–1·31)	0·10
Obese: BMI >30 kg/m^2 (%)	29·6 (26·5–32·8)	25·4 (22·6–28·5)	28·4 (25·4–31·6)	22·6 (20·0–25·6)	0·86 (0·81–0·94)	<0·0001
Daily alcohol consumption (%)	16·3 (13·9–18·9)	15·6 (13·3–18·3)	14·1 (11·9–16·6)	20·8 (18·2–23·7)	1·15 (1·06–1·25)	0·001
Childhood environment						
Adult height (mm)	1591·0 (1586·9–1595·1)	1588·4 (1584·3–1592·5)	1587·9 (1583·8–1591·9)	1584·9 (1580·9–1589·0)	1·45 (0·91–2·05)‡	<0·0001
Adult leg length (mm)	757·8 (755·0–760·6)	756·0 (753·2–758·8)	756·5 (753·2–758·8)	758·6 (755·8–761·4)	1·19 (0·67–1·76)‡	0·01
Adult leg to trunk ratio %	91·0 (90·6–91·3)	91·3 (90·9–91·6)	91·2 (90·7–91·5)	91·8 (91·4–92·1)	0·32 (0·14–0·50)‡	0·001

Data are prevalence (95% CI) for socioeconomic indicators and behavioural and lifestyle risk factors, and means (95% CI) for biomarkers childhood environment. All estimates are adjusted for age and month of blood sampling (so seasonal effects are controlled for). *Per increase in 1 SD vitamin E. †Ex smoker=compared with never smokers, does not include current smokers. ‡Difference per increase in 1 SD vitamin E.

Table 2: **Life course factors across quartiles of plasma vitamin E**

general, non-differential measurement error of exposures will lead to underestimates of the exposure-outcome measures. However, in multivariable analyses, if covariates are measured with different levels of precision, the adjusted relative risk can be biased in either direction.[14] Another important issue, which is often not addressed in observational studies, is that of ensuring that covariates are modelled correctly, taking into account non-linear associations and interactions. Taking repeated measurements on a smaller sample will often be the best strategy in situations of appreciable confounding.[14]

A life course approach

In an ideal world, prospective transgenerational observational studies would establish accurate measures of exposures from before conception to death. In the real world such studies will be many years in the making. However, it is possible, as in the example presented here, to ask about early life exposures and also to measures adult biomarkers—such as stature and in particular leg length—of early life exposures.

The probability that residual confounding explains any remaining association when attenuation for measured confounders occurs in a study should always be considered. For example, in one study the relative risk of myocardial infarction associated with a low vitamin C concentration was reduced from 4·03 (1·74–9·36) to 2·08 (0·82–5·30) after adjustment for a range of confounders.[15] Similarly, an association between baseline consumption of dietary fruit and vegetables and future risk of lung cancer was greatly attenuated by adjustment for several adult risk factors, including smoking.[16] In both these studies the investigators suggested that their results were consistent with beneficial effects of antioxidant vitamins. In both cases the remaining apparent effect is in our view, probably explained by residual confounding. Measurement error of the confounders included in the adjusted model, and also failure to include relevant confounding factors from across the life course, will be involved.[17]

Sensitivity analyses to assess residual confounding

Sensitivity analyses can be undertaken to give an indication of the probable effect of measurement error for included confounders, and of unmeasured confounders; methods have been described in detail elsewhere.[13] For unmeasured confounders various plausible values for the strength of associations between the unmeasured confounder and outcome; and of unmeasured confounder with exposure, can be used, and a series of fully adjusted exposure-outcome associations can be estimated. If these estimated associations are little attenuated, unmeasured confounders are unlikely to be a major issue. An alternative way to present these sensitivity analyses is to report the strength of an association between all potential unmeasured confounders with exposure and outcome that would be needed to attenuate an association to the null value.

Specificity of an association

Specificity refers to the idea that an association is more likely to be causal if an exposure is related to one outcome, and was one of the Bradford Hill's indicators of causality. It has lost favour over recent decades: for example, a standard textbook of epidemiology states that "Specificity does not confer greater validity to any causal influence regarding the exposure effect . . . the criterion is useless and misleading."[18] In support of this, it is pointed out that smoking is related to many causes of death, but that does not exonerate it as a cause of lung cancer.[18] Smoking, however, is not a single exposure—tobacco smoke contains a wide variety of factors, each of which might be specifically associated with different outcomes. Smoking is unusual in this regard. The view that specificity is an

unimportant criterion for causation has been challenged, since in many cases an exposure is a plausible cause of just one or a few outcomes; observing associations with both the hypothesised outcome or outcomes and with other outcomes decreases the probability of the former being causal.[19] For example, in 1986, Petitti and colleagues[20] showed that hormone replacement therapy was apparently "protective" against deaths from accidents and violence (as well as cardiovascular disease) in an observational study, a finding resistant to statistical adjustments. This suggested that both hormone replacement therapy-outcome associations were residually confounded. This report has been rarely cited and failed to quell the interest in promoting hormone replacement therapy as a preventive intervention for cardiovascular disease risk. In the EPIC cohort study high plasma vitamin C concentrations were associated with reductions in all-cause, cardiovascular, and cancer mortality.[1] Again, this suggests that confounding by a wide range of socioeconomic and behavioural factors is involved. Examination of associations of high concentrations of antioxidants with risk of deaths from accidents and violence, or other implausibly causal outcomes would be informative.

Mendelian randomisation

Functional genetic polymorphisms that mimic the effects of environmental exposures can be used to provide more robust tests of relations between exposure and disease. There is little opportunity when alleles segregate—effectively a random process—for social and behavioural factors to confound the resulting association between the polymorphism and disease.[21] As genomic and proteomic sciences advance, observational studies of allelic associations might provide robust evidence of causal associations between environmentally modifiable exposures and diseases.[21]

Conclusions

The disparity between observational studies and randomised trial evidence of the health effects of antioxidant vitamins is probably explained by a failure to appreciate the complex and important differences between adults with high vitamin concentrations and those with lower concentrations. High intake of antioxidant vitamins might not be causally related to cardiovascular and other diseases, but rather serves as a proxy indicator of a host of factors that protect against these diseases. Randomised controlled trials provide the most robust estimate of causal effect when they are feasible. However, they are not always feasible and because of the expense and ethical concerns of randomised trials, it is important that observational studies are used to effectively direct investigators to the interventions most appropriately assessed by trials. Abandoning observational studies for randomised controlled trials would not therefore be a panacea. A more careful approach to the design and analysis of observational epidemiological studies should ensure that they remain a useful method for generating and testing hypotheses which ultimately might improve the health of the public.

Contributors

G Davey Smith, S Ebrahim, and D A Lawlor developed the idea for this piece. S Ebrahim and D A Lawlor were involved in data collection on the British Women's Heart and Health Study. D Kundu and K R Bruckdorfer undertook the vitamin assays. D A Lawlor and S Ebrahim undertook the analysis included in this report. D A Lawlor wrote the first draft and coordinated writing of the report. All authors contributed to the final version. D A Lawlor acts as guarantor for the paper

Conflict of interest statement

None declared.

Acknowledgments

The British Women's Heart and Health Study is funded by the Department of Health. DAL was funded by a Medical Research Council/Department of Health training fellowship when this work was done and is now funded by a Department of Health Career Scientist Award. P Whincup helped obtain funds from the British Heart Foundation for the vitamin assays used in this study. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health or Medical Research Council.

References

1 Khaw KT, Bingham S, Welch A, 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.

2 Rimm EB, Stampfer MJ, Ascherio A, Giovannuci E, Colditz GA, Willet WC. Vitamin E consumption and the risk of coronary heart disease in men. *N Engl J Med* 1993; **328:** 1450–56.

3 Stampfer MJ, Hennekens CH, Manson JE, Colditz GA, Rosner B, Willett WC. Vitamin E consumption and the risk of coronary heart disease in women. *N Engl J Med* 1993; **328:** 1444–49.

4 The Alpha-Tocopherol Beta Carotene cancer prevention study group. The effect of vitamin E and Beta carotene on the incidence of lung cancer and other causes in male smokers. *N Engl J Med* 1994; **330:** 1029–35.

5 Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* 2002; **360:** 23–33.

6 Steinberg D, Witztum JL. Is the oxidative modification hypothesis relevant to human atherosclerosis? Do the antioxidant trials conducted to date refute the hypothesis? *Circulation* 2002; **105:** 2107–11.

7 Collaborative Group of the Primary Prevention Project. Low-dose aspirin and vitamin E in people at cardiovascular risk: a randomised trial in general practice. *Lancet* 2001; **357:** 89–95.

8 Hennekens CH, Buring JE, Manson JE, et al. Lack of effect of long-term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease. *N Engl J Med* 1996; **334:** 1145–49.

9 Davey Smith G. Reflections on the limitations to epidemiology. *J Clin Epidemiol* 2001; **54:** 325–31.

10 Gunnell D. Can adult anthropometry be used as a biomarker for prenatal and childhood exposures? *Int J Epidemiol* 2002; **31:** 390–94.

11 Lawlor DA, Ebrahim S, Davey Smith G. The association between components of adult height and type II diabetes and insulin resistance: British Women's Heart and Health Study. *Diabetologia* 2002; **45:** 1097–106.

12 Davey Smith G, Hart C. Lifecourse socioeconomic and behavioural influences on cardiovascular disease mortality: the Collaborative Study. *Am J Public Health* 2002; **92:** 1295–98.

13 Greenland S. Basic methods for sensitivity analysis of biases. *Int J Epidemiol* 1996; **25:** 1107–16.

14 Phillips AN, Davey Smith G. The design of prospective epidemiological studies: more subjects or better measurements? *J Clin Epidemiol* 1993; **46:** 1203–11.

15 Nyyssonen K, Parviainen MT, Salonen R, Tuomilehto J, Salonen JT. Vitamin C deficiency and risk of myocardial infarction: prospective population study of men from eastern Finland. *BMJ* 1997; **314:** 634–38.

16 Holick CN, Michaud DS, Stolzenberg-Solomon R, et al. Dietary carotenoids, serum beta-carotene, and retinol and risk of lung cancer in the alpha-tocopherol, beta-carotene cohort study. *Am J Epidemiol* 2002; **156:** 536–47.

17 Phillips AN, Davey Smith G. How independent are "Independent" effects? Relative risk estimation when correlated exposures are measured imprecisely. *J Clin Epidemiol* 1991; **44:** 1223–31.

18 Rothman KJ, Greenland S. Modern epidemiology, 2nd edn. Philadelphia: Lippincott Williams and Wilkins, 1998: 25.

19 Weiss NS. Can the "specificity" of an association be rehabilitated as a basis for supporting a causal hypothesis? *Epidemiology* 2002; **13:** 6–8.

20 Petitti DB, Perlman JA, Sidney S. Postmenopausal estrogen use and heart disease. *N Engl J Med* 1986; **315:** 131–32.

21 Davey Smith G, Ebrahim S. "Mendelian randomisation": can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; **32:** 1–22.