Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Subgroup analysis of large trials can guide further research: a case study of vitamin E and pneumonia

This article was published in the following Dove Press journal:
Clinical Epidemiology
14 February 2011
Number of times this article has been viewed

Harri Hemilä
Jaakko Kaprio

Department of Public Health, University of Helsinki, Helsinki, Finland

Background: Biology is complex and the effects of many interventions may vary between population groups. Subgroup analysis can give estimates for specific populations, but trials are usually too small for such analyses.

Purpose: To test whether the effect of vitamin E on pneumonia risk is uniform over subgroups defined by smoking and exercise.

Methods: The Alpha-Tocopherol Beta-Carotene Cancer Prevention Study examined the effects of vitamin E (50 mg per day) and β-carotene (20 mg per day) on lung cancer in 29,133 male smokers aged 50–69 years using a 2 × 2 factorial design. The trial was conducted among the general community in Finland during 1985–1993; the intervention lasted for 6.0 years (median). In the present study, we tested the uniformity of vitamin E effect on the risk of hospital-treated pneumonia (898 cases) by adding a dummy variable to allow each subgroup its own vitamin E effect in a Cox model covering all participants.

Results: Vitamin E effect was not uniform over eight subgroups defined by baseline smoking (5–19 vs ≥20 cigarettes per day), age of smoking initiation (≤20 vs ≥21 years), and exercise during leisure time (yes vs no). Vitamin E decreased pneumonia risk by 69% (95% CI: 43% to 83%) among participants who had the least exposure to smoking and exercised during leisure time. Vitamin E increased pneumonia risk by 79% (95% CI: 27% to 150%) among those who had the highest exposure to smoking and did not exercise.

Limitations: Although the evidence of heterogeneity is strong, it is not evident to what extent the estimates of effect or the limits between the subgroups can be extrapolated to other populations.

Conclusion: Subgroup analysis of large trials should be encouraged, though caution is needed in the interpretation of findings. The role of vitamin E in susceptibility to pneumonia in physically active nonsmokers warrants further study.

Trial registration: ClinicalTrials.gov NCT00342992.

Keywords: vitamin E, pneumonia, smoking, leisure time exercise, α-tocopherol, β-carotene, subgroup analysis

Introduction

The size of a controlled trial is usually based on a power calculation, the goal of which is to determine the minimal number of participants needed to test whether an overall difference exists between the intervention and control groups. Such trials are too small to test subgroup differences. Furthermore, carrying out numerous subgroup comparisons leads to the multiple testing problem. Such reasoning is the major cause for discouraging subgroup analyses.[1–5]

Correspondence: Harri Hemilä
Department of Public Health, University of Helsinki, POB 41, Helsinki, FIN-00014, Finland
Tel +358 9 191 27573
Fax +358 9 191 27570
Email harri.hemila@helsinki.fi

The above argument has limitations, however. For example, if a trial collects data on a secondary outcome which are much more numerous than the primary outcome, say lung cancer, subgroup analysis on the secondary outcome, such as the common cold,[6] does not suffer from low statistical power. Furthermore, most controlled trials study the effect of drugs having a specific biochemical target within patients who are narrowly selected, and a large within-trial variation in the effect may be unlikely in such cases. However, it is possible that the within-trial variation in the effect is substantially greater for interventions that have complex and broad effects on the human system, in particular when the effects are studied in heterogeneous populations. Thus, while reasons exist for being cautious about subgroup analysis in general, there are conditions when subgroup analyses may be justified.

Previously, we explored the effect of vitamin E on pneumonia risk among the 29,133 male smokers of the Alpha-Tocopherol Beta-Carotene [ATBC] Study.[7,8] We found significant modification of vitamin E effect by age of smoking initiation, in that the vitamin reduced the risk in those who started smoking at a late age and, within this subgroup, baseline smoking further modified the effect so that the benefit was greatest among those who smoked the least.[9] Since physical activity leads to oxidative stress,[10] we separately hypothesized that vitamin E might reduce pneumonia risk among physically active ATBC Study participants, and found that the vitamin halved the risk in those who exercised during leisure time.[11] These findings indicate that cigarette smoking and exercise might modify the effect of vitamin E on pneumonia risk. However, since several comparisons were made, the multiple testing problem cannot be entirely dismissed. Therefore, in this paper we analyze the subgroup differences in all ATBC Study participants simultaneously.

If there is firm evidence that the effect of vitamin E supplementation on health outcomes of the ATBC participants is heterogeneous, this would imply that subgroup analyses in other large-scale trials on vitamin E, and possibly in large-scale trials on other subjects, should be encouraged rather than discouraged.

Material and methods

Participants

The rationale, design, and methods of the ATBC Study examining the effects of vitamin E (*dl*-α-tocopheryl acetate, AT, 50 mg/day) and β-carotene (BC, 20 mg/day) on the incidence of lung cancer and other cancers have been described in detail.[7–9] The ATBC Study is registered at ClinicalTrials.gov under the identifier NCT00342992.

In brief, to be eligible, male participants aged 50–69 years had to smoke ≥5 cigarettes per day at entry, and those enrolled in the trial (N = 29,133) were randomized to one of four intervention arms and administered placebo, AT, BC, or AT+BC, using a 2 × 2 factorial design. Compared with baseline levels, supplementation increased the serum level of α-tocopherol by 50%.[7,8] The intervention continued for 5 to 8 years until April 1993. The trial was approved by the review boards of the participating institutions and all participants gave written informed consent. Compliance with supplementation was high: some 90% of the subjects took more than 90% of their prescribed capsules during their active participation in the trial.[7,8]

Baseline characteristics

Before randomization at baseline, the participants completed questionnaires on medical and smoking histories and general background characteristics. A detailed dietary history questionnaire was completed that provided data regarding vitamins C and E, and coffee consumption.[12] Age of smoking initiation was not available for seven participants and dietary data for 2,022 participants.

Previously, we found that dichotomization of the age of smoking initiation with the cutoff point at 21 years appropriately captured the variation of the vitamin E effect,[9] and the same cutoff was used in this study. Although smoking is a continuous variable, it is heavily clustered to multiples of 20 (and 10) cigarettes per day. In this study, we dichotomized cigarette smoking to 5–19 cigarettes per day and to ≥20 per day. As we recognized that in both cases dichotomization leads to a loss of information of the continuous variables, we examined the effect of vitamin E in smaller ranges in Tables 2 and 3.

The baseline questionnaire on physical activity during leisure time was a modification of that used originally in the Gothenburg study focusing on cardiovascular diseases.[13] The intensity of average physical activity during leisure time over the previous 12 months was enquired about using the following alternatives: 1) light: reading, watching TV, listening to the radio, or going to movies, ie, activities that are not physically demanding; 2) moderate: walking, fishing, hunting, or gardening quite regularly; and 3) heavy: actual physical exercise, such as jogging, skiing, swimming, gymnastics, and court and field sports quite regularly. In the current analyses we combined answers 2) [n = 15,191] and 3) [n = 1,744] to the category "exercise during leisure time". Data on exercise were not available for 14 participants.

Outcome and follow-up time

The events for this study, the first hospital-treated cases of pneumonia after randomization, were ascertained from the national Hospital Discharge Register using the unique personal identification numbers for linkage (see details in Hemilä et al)[9]. Pneumonia cases recorded in the Hospital Discharge Register reflect clinically more severe cases of greater health and economic significance, whereas less severe cases of pneumonia treated as outpatients are not recorded in the Register. Use of the Hospital Discharge Register allowed for the obtaining of information on pneumonia in all study participants irrespective of whether they continued in or had dropped out of the trial.

Follow-up time for each participant began from the day of randomization, and continued until the date of first hospital discharge for pneumonia, death, or the end of the trial, April 30, 1993, whichever came first. The median follow-up time of the participants was 6.0 years, and there was a total of 167,968 person-years of observation.

Statistical methods

We estimated the effect of vitamin E supplementation on pneumonia incidence through Cox models. We calculated the risk ratio (RR) and the 95% confidence interval (CI) of the RR using the PROC PHREG program of the SAS package of programs (release 8.2, SAS Institute, Inc., Cary, NC). No covariates were included in the models analyzing the treatment effects. As to supplementation, we carried out the analyses following the intention-to-treat (ITT) principle.

In Table 1, we compared the trial participants administered vitamin E (AT and AT + BC) with those not receiving vitamin E (the no-vitamin E participants; placebo and BC). Since, in Table 3, we observed that AT and BC supplementations interacted, we restricted further subgroup analyses of Table 3 to the no-BC participants (AT and placebo arms). Because of this interaction, we also re-tested the heterogeneity of Table 1 by restricting to the no-BC participants.

Table 1 The effect of vitamin E on pneumonia incidence by level of cigarette smoke exposure and exercise during leisure time: ATBC Study 1985–1993

Age of smoking initiation (years)	Cigarettes per day at baseline		Effect of vitamin E	
			Exercise during leisure time	
			Yes	No
≥21	5–19	RR[a]	0.31	0.85
		(95% CI)[a]	(0.17, 0.57)	(0.44, 1.64)
		Cases of pneumonia[b]	14/43	17/19
		No. of men[c]	2,216	1,043
≥21	≥20	RR[a]	0.84	0.86
		(95% CI)[a]	(0.48, 1.46)	(0.50, 1.49)
		Cases of pneumonia[b]	24/27	24/28
		No. of men[c]	2,445	1,763
≤20	5–19	RR[a]	1.24	1.05
		(95% CI)[a]	(0.87, 1.78)	(0.71, 1.56)
		Cases of pneumonia[b]	68/56	51/50
		No. of men[c]	4,602	2,688
≤20	≥20	RR[a]	0.88	1.35
		(95% CI)[a]	(0.67, 1.15)	(1.06, 1.73)
		Cases of pneumonia[b]	97/110	152/115
		No. of men[c]	7,669	6,686

Notes: [a]The Cox model comparing participants who received vitamin E with those who did not; [b]A/B refers to A cases of pneumonia among the vitamin E participants and B cases of pneumonia among the no-vitamin E participants. Data on age of smoking initiation or exercise were missing from two pneumonia cases among the vitamin E participants and from one case among the no-vitamin E participants; these cases are not included in this table; [c]The number of participants in the vitamin E and no-vitamin E groups was the same within 5% accuracy in each of the eight groups. The uniformity of the vitamin E effect was tested by adding a dummy variable for vitamin E effect in seven groups of the table, allowing each of the eight groups their own vitamin E effect. The regression model was improved by $\chi^2(7\ df) = 26.6$, $P = 0.0004$, compared to the model with a uniform vitamin E effect. Heterogeneity is mainly caused by the upper-left and lower-right cells: the addition of only these two cells improved the model by $\chi^2(2\ df) = 23.4$. The difference between the above two models is fully explained by chance: $\chi^2(5\ df) = 3.2$. The addition of the third-order interaction term, between vitamin E supplementation, age of smoking initiation, cigarettes per day, and leisure time exercise, to the model containing all lower level interaction terms, improved the regression model by $\chi^2(1\ df) = 5.4$, $P = 0.02$. Since vitamin E and β-carotene supplementations interact in the lower-right cell (see Table 3), we also tested the uniformity of vitamin E effect among the no-β-carotene participants (n = 14,564). Adding a dummy variable for vitamin E effect in seven groups of the table improved the model by $\chi^2(7\ df) = 22.8$, $P = 0.002$. Adding only the upper-left and lower-right cells improved the model by $\chi^2(2\ df) = 17.8$, indicating that the effect of vitamin E is restricted to the upper-left and lower-right cells. The difference between the two models is fully explained by chance: $\chi^2(5\ df) = 5.0$. Nevertheless, adding the third-order interaction term to a model containing all lower level interactions did not significantly improve the model: $\chi^2(1\ df) = 2.0$, $P = 0.16$. Vitamin E and β-carotene supplementations did not interact in cells of this table other than the lower-right cell.

Abbreviations: RR, risk ratio; CI, confidence interval.

To test the statistical significance of interaction between vitamin E supplementation and potential modifying factors, we first added vitamin E and the modifying factor to the regression model. The statistical significance of the interaction was thereafter calculated from the change in $-2 \times \log$ (likelihood) when the interaction term for vitamin E supplementation and the modifying factor were added to the model. In our subgroup analyses in Tables 2 and 3, we split the subgroup variables at levels leading to a reasonably similar number of cases in the control groups.

Nelson-Aalen cumulative hazard functions were constructed using the STATA sts program (Release 9, Stata Corp, College Station, TX). Two-tailed *P*-values are presented.

Results

Among all ATBC participants, the cases of pneumonia were identically divided between the vitamin E and no-vitamin E groups: 449 vs 449, corresponding to RR = 1.00 (95% CI: 0.88, 1.14).

We divided the participants into eight subgroups on the basis of age of smoking initiation, level of smoking at the baseline of the trial, and exercise during leisure time (Table 1). We tested the uniformity of the vitamin E effect by adding a dummy variable for vitamin E effect in seven groups of the table, and this significantly improved the Cox model ($P = 0.0004$). The heterogeneity in Table 1 is fully explained by the upper-left and lower-right corners, ie, by the opposite corners of the table. Furthermore, the third-level interaction term between vitamin E supplementation, age of smoking initiation, level of smoking, and exercise was significant when comparing the vitamin E and no-vitamin E participants. Since the effect of vitamin E was restricted to the upper-right and lower-left corners, we analyzed these two groups further.

Among the 2,216 participants who initiated smoking at a late age, smoked less than a pack of cigarettes per day, and exercised during leisure time, vitamin E supplementation reduced pneumonia risk by 69% (upper-left cell in Table 1; Figure 1). The estimated effect of vitamin E in this subgroup

Table 2 The effect of vitamin E on pneumonia incidence in ATBC participants who initiated smoking at ≥21 years, smoked 5–19 cigarettes per day, and exercised during leisure time

Subgroup	No. of men[a]	Cases of pneumonia[b]	Effect of vitamin E	
			RR (95% CI)[c]	Test for interaction (*P*)
All	2,216	14/43	0.31 (0.17, 0.57)	
β-Carotene supplementation				
No	1,093	6/19	0.32 (0.13, 0.80)	0.9
Yes	1,123	8/24	0.30 (0.13, 0.66)	
Cigarettes (1/day)				
5–11	1,110	6/18	0.31 (0.12, 0.78)	1.0
12–19	1,106	8/25	0.31 (0.14, 0.69)	
Age of smoking initiation (years)				
21–25	1,563	9/28	0.30 (0.14, 0.63)	0.9
26–60	653	5/15	0.34 (0.12, 0.92)	
Age at baseline (years)				
50–64	1,810	8/23	0.33 (0.14, 0.73)	0.9
65–69	406	6/20	0.30 (0.12, 0.75)	
Dietary vitamin E (mg/day)[d]				
<9	570	5/20	0.26 (0.10, 0.69)	0.6
≥9	1,486	8/20	0.36 (0.16, 0.81)	
Dietary vitamin C (mg/day)[d]				
<100	1,030	8/20	0.37 (0.16, 0.85)	0.5
≥100	1,026	5/20	0.23 (0.09, 0.62)	
Coffee (mL/day)[d]				
<500	999	9/18	0.47 (0.21, 1.05)	0.12
≥500	1,057	4/22	0.17 (0.06, 0.48)	

Notes: [a]The number of participants in the vitamin E and no-vitamin E groups was the same within 8% accuracy in all subgroups shown; [b]A/B refers to A cases of pneumonia among the vitamin E participants and B cases of pneumonia among the no-vitamin E participants; [c]The Cox model comparing participants who received vitamin E with those who did not; [d]Data on diet were missing for 160 participants, which included one case of pneumonia in the vitamin E group and three cases in the no-vitamin E group.
Abbreviations: RR, risk ratio; CI, confidence interval.

Table 3 The effect of vitamin E on pneumonia incidence in ATBC participants who initiated smoking at ≤20 years, smoked ≥20 cigarettes per day, and did not exercise during leisure time

Subgroup	No. of men[a]	Cases of pneumonia[b]	Effect of vitamin E	
			RR (95% CI)[c]	Test for interaction (*P*)
All	6,686	152/115	1.35 (1.06, 1.7)	
β-Carotene supplementation				
No	3,371	89/51	1.79 (1.27, 2.5)	0.02
Yes	3,315	63/64	1.01 (0.71, 1.4)	
Restriction to the no-β-carotene participants:				
No β-carotene	3,371	89/51	1.79 (1.27, 2.5)	
Cigarettes (1/day)				
20–25	2,269	62/36	1.78 (1.18, 2.7)	1.0
26–80	1,102	27/15	1.83 (0.97, 3.5)	
Age of smoking initiation (years)				
6–17	1,616	48/26	1.94 (1.20, 3.1)	0.6
18–20	1,755	41/25	1.64 (1.00, 2.7)	
Age at baseline (years)				
50–59	2,466	55/31	1.84 (1.19, 2.9)	0.8
60–69	905	34/20	1.70 (0.98, 3.0)	
Dietary vitamin E (mg/day)[d]				
<9	1,231	31/22	1.52 (0.88, 2.6)	0.5
≥9	1,909	49/26	1.90 (1.18, 3.1)	
Dietary vitamin C (mg/day)[d]				
<70	1,229	38/22	1.76 (1.04, 3.0)	0.9
≥70	1,911	42/26	1.69 (1.03, 2.8)	
Coffee (mL/day)[d]				
<500	1,188	38/20	1.95 (1.13, 3.4)	0.5
≥500	1,952	42/28	1.56 (0.96, 2.5)	

Notes: [a]The number of participants in the vitamin E and no-vitamin E groups was the same within 5% accuracy in all subgroups shown; [b]A/B refers to A cases of pneumonia among the vitamin E participants and B cases of pneumonia among the no-vitamin E participants; [c]The Cox model comparing participants who received vitamin E with those who did not; [d]Data on diet were missing for 231 participants, which included nine cases of pneumonia in the vitamin E group and three cases in the no-vitamin E group.
Abbreviations: RR, risk ratio; CI, confidence interval.

was robust in several further subgroup analyses. The effect was not modified by BC supplementation, age, or dietary vitamins C and E (Table 2). Dividing the participants by the age of smoking initiation and baseline smoking also led to compatible effects within the smaller subgroups. Previously, we found that coffee consumption significantly modified the benefit of vitamin E in those who started smoking at a late age.[9] The subgroup differences in Table 2 are in line with the earlier findings, but not significantly.

Among the 6,686 participants who initiated smoking at an early age, smoked a pack of cigarettes daily or more, and did not exercise, vitamin E increased pneumonia risk by 35% when compared with the no-vitamin E group (lower-right cell in Table 1). However, in this subgroup the vitamin E effect was modified by BC supplementation so that the harm of vitamin E was restricted to those who were not administered BC (Table 3). Therefore, we restricted the further subgroup analyses of Table 3 to the no-BC participants. Among the no-BC participants, vitamin E increased pneumonia risk by 79%, and this effect was robust in further subgroup analyses (Table 3).

Previously, we hypothesized that the marginally significant 14% increase in pneumonia risk among those ATBC participants who started smoking at an early age (n = 21,657; the four lowest cells in Table 1) might correspond to a more unambiguous harmful effect among low-weight participants, based on an assumption of dose-dependency.[14] Then we found that vitamin E increased pneumonia risk in participants weighing less than 60 kg. Unexpectedly, vitamin E also increased pneumonia risk at the opposite end of the weight scale, among those weighing over 100 kg.[14] Furthermore, in both groups, harm caused by vitamin E was restricted to those who had a dietary vitamin C intake above the median. Therefore, we examined whether weight and vitamin C intake might modify the effect of supplementation outside of the lower-right corner in Table 1.

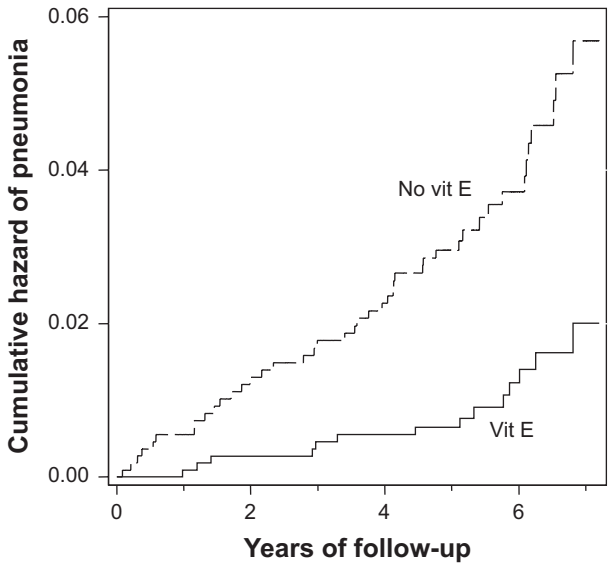

Figure 1 Vitamin E and pneumonia risk in ATBC Study participants who started smoking at ≥21 years, smoked 5–19 cigarettes per day, and exercised (n = 2,216). Nelson-Aalen cumulative hazard functions for vitamin E and no-vitamin E groups are shown. Each step indicates one case of pneumonia. For the difference between the two survival curves, the logrank test gives $P = 0.00005$. The survival curves are cut at 7.2 years because the number of participants declines abruptly thereafter (no cases after 6.8 years). At six-year follow-up 576 and 535 participants remained in the vitamin E and the no-vitamin E groups, respectively.

Of the low-weight high vitamin C participants, 72% (337 of 468) were outside the lower-right corner of Table 1; in these 337 participants there were 19 pneumonia cases among the vitamin E and eight cases among the no-vitamin E participants (RR = 2.7, 95% CI: 1.18–6.2). Of the overweight high vitamin C participants, 65% (397 of 613) were outside the lower-right corner of Table 1; in these 397 participants there were 10 pneumonia cases among the vitamin E and one case among the no-vitamin E participants ($P = 0.01$, Fisher's test). Consequently, weight and dietary vitamin C appear to modify the effect of vitamin E independent of smoking and exercise.

Discussion

The numbers of pneumonia cases in the ATBC Study were equally distributed between the vitamin E and no-vitamin E participants, indicating a lack of overall effect with great accuracy. However, in this study we have shown that the effect of vitamin E is not uniformly nil over all the ATBC Study population. Depending simultaneously on the two different measures of cigarette smoking and on the level of exercise, vitamin E supplementation decreased, increased or had no effect on the incidence of pneumonia (Table 1).

Among those who had the least exposure to smoking and exercised during leisure time, vitamin E decreased the risk of pneumonia by 69%. This group covers 8% of the ATBC Study participants. The effect estimate was robust in further subgroup analyses (Table 2).

The group that had the highest exposure to smoking and did not exercise covered 23% of the ATBC participants. In this group, vitamin E increased pneumonia risk by 79% in the no-BC participants (Table 3). This effect estimate was also robust in further subgroup analyses, however simultaneous BC supplementation nullified the harmful effects of vitamin E.

In our subgroup analysis focusing on smoking and exercise, 69% of the ATBC participants fell into the six middle groups that were consistent with vitamin E having no effect (Table 1). Nevertheless, it is possible that there are further modifying factors in addition to smoking and physical activity. Previously, we found that coffee drinking modified the effect of vitamin E among those who started smoking at a late age.[9] Among those who started smoking at an early age, weight and dietary vitamin C intake modified the vitamin E supplementation effect.[14] The current analyses are not inconsistent with these earlier subgroup findings. Thus, it seems possible that vitamin E can affect pneumonia risk in some groups of people depending on six or more modifying factors meaning that the modification is complex and does not follow a simple multiplicative model.

It is often suggested that subgroup findings should be trusted only when they are replicated in other trials. Although such a suggestion seems sound, the heterogeneity we found in the effect of vitamin E on pneumonia suggests that testing a subgroup difference in another sample of people can be all but simple. When the effect of vitamin E may depend simultaneously on six or more modifying factors, the findings for the first-level interactions depend on the selection of participants.

For example, in the whole ATBC Study, baseline smoking did not modify the effect of vitamin E ($P = 0.2$).[9] However, Table 1 indicates that baseline smoking modifies the vitamin E effect conditionally on the age of smoking initiation and the level of exercise. This means that depending on the composition of the population, baseline smoking may or may not modify the effect of vitamin E. Similarly, we previously found that vitamin E halved the risk of pneumonia in ATBC participants who exercised during leisure time;[11] however, Table 1 indicates that this effect is conditional on low level of exposure to smoking. On the basis of these examples, replication is not a universally valid method for deciding whether the subgroup differences observed in one trial are real or not.

Peto et al argued that "believing that a treatment effect exists in one stratum of patients, even though no overall significant treatment effect exists, is a common error".[4] This comment may be sound with respect to rather small therapeutic trials. However, Table 1 and our previous ATBC Study subgroup analyses[6,9,11,14–17] show that there can be strong evidence of vitamin E effect in specific groups of people, even though no overall effect exists. Accordingly, Peto et al's argument should not be taken as a universal objection to analyzing subgroups in the absence of overall effect.

Several investigators have strongly discouraged subgroup analysis.[1–5] However, other authors have considered that a universal denial of subgroup analysis is an exaggerated reaction. Feinstein wanted to "rescue the scientific importance of valid pathophysiologic subgroups from being forgotten or destroyed by excessive vehemence in suggestions that all subgroups are evil".[18] Lagakos noted that "avoiding any presentation of subgroup analysis because of their history of being overinterpreted is a steep price to pay for a problem that can be remedied by more responsible analysis and reporting".[19] Rothwell responded to popular arguments against subgroup analysis and described situations where subgroup analysis seems to be justified.[20]

Altman considered that biological plausibility is a weak criterion when deciding whether a subgroup finding is likely to be real, since, according to him, physicians seem able to find a biologically plausible explanation for any finding.[2] There is much room for speculation at the biochemical level, because the number of genes and their effects is huge, and Altman's argument can have validity in such a context. However, the number of variables relevant at the population level of biology is much more limited. For example, few factors compare with the importance of smoking as a factor influencing the health of the lungs. Physical activity is also a fundamentally important factor determining health. Smoking affects the metabolism of vitamin E[21] and sporadic physical stress causes oxidative stress which is not compensated by an increase in antioxidative enzyme levels, unlike regular physical activity.[10] Therefore, both smoking and exercise are plausible modifying factors for the effects of vitamin E supplementation, which increases the credibility of the heterogeneity seen in Table 1.

Previously, two small trials examined the effect of vitamin E on respiratory infections in elderly people, both with less than 700 participants and lasting for about one year. In the first, Meydani et al calculated 13 *P*-values for ITT comparisons between 200 mg/day vitamin E and placebo groups, and only one of them suggested that vitamin E might reduce the incidence of respiratory infections, yet very marginally so ($P = 0.048$).[22] In the second, Graat et al found that 200 mg/day of vitamin E did not influence the incidence of respiratory infections, yet made the symptoms more severe ($P = 0.02$).[23] Because both of these trials are small and there are differences in outcome definitions etc, it is not possible to decide whether their findings are inconsistent or not. Graat et al's findings indicating harmful effects of vitamin E conflict with the wide spread belief that the vitamin is beneficial, or at least not harmful.[24] Therefore, it is not obvious whether Graat et al's findings should be interpreted as a reflection of real harm or as a result of chance. Given the strong evidence of heterogeneity we observed in the effect of vitamin E on pneumonia (Table 1) and on the common cold,[6] it seems plausible that the harmful effects observed by Graat et al are real and are explained by the selection of participants, but do not reflect a universal harmful effect of vitamin E. In this respect, the observed heterogeneity in the ATBC Study can influence the interpretation of smaller trials. Nevertheless, we are skeptical as regards the possibility of extrapolating the effect estimates and the exact limits of the subgroups of Table 1 to other contexts.

Although the division of participants on the basis of baseline physical activity and smoking is sound, both of these factors can change with time. Some participants stopped exercising or smoking over the several-years-long follow-up, yet they remained classified in the same subgroups. This phenomenon can dilute the differences between the subgroups and shift the estimates of effect closer to unity; however, it cannot explain the significant heterogeneity observed when the participants are divided by the baseline measurements. Furthermore, exercise and smoking are correlated with numerous other life style variables and we cannot dismiss the possibility that other life style factors might be behind the heterogeneity observed in Table 1. Nevertheless, this concern does not challenge the evidence indicating that substantial heterogeneity exists across various population groups in the effect of vitamin E on pneumonia risk, even if the real modifying variables might be different from those used for defining the subgroups of Table 1.

The ATBC Study included 29,133 participants which is over 40 times more than the number of participants in the Meydani et al[22] and Graat et al[23] trials. In this respect, a large trial can be considered as a series of smaller trials when there is sound justification for setting the borders between the subgroups. A particular strength of a subgroup analysis of

a large trial is that the intervention and outcome definitions are identical over the trial. Therefore, subgroup analysis of a large trial can yield much more valid explanations for the heterogeneity of effect compared with the analysis of the heterogeneity of small trials that have numerous concurrent differences.

For many diseases, recognized risk factors account for at best only a modest fraction of variation in disease risk. Much effort is put into identifying new factors, either environmental or genetic. Our analyses indicate that complex patterns of interaction, perhaps in a context-specific manner, may also contribute to disease risk. Such effects may thus account for some of the unexplained variability of disease risk.

Our subgroup analyses of the respiratory infections of ATBC participants[6,9,14,15] made it also possible to hypothesize that the identified modifying factors might modify the effect of vitamin E on the mortality of these participants. We found that, conditional on a high level of dietary vitamin C intake, age modified the effect of vitamin E on mortality.[16,17] Thus, we could partially extrapolate the modifying factors identified in the subgroup analyses on respiratory infections to an outcome that has a very weak relation to such infections.

Vandenbroucke pointed out that medical science has two divergent goals.[25] First, controlled trials test whether an intervention works or not. Second, most basic medical science emphasizes discovery – searching for the biological mechanisms and causes of diseases, and for explanations in general. This divergence in views is relevant when considering a proper attitude to subgroup analysis. Evidently, great caution must be exercised when proposing a treatment on the basis of unanticipated subgroup findings. On the other hand, subgroup analysis can generate new hypotheses and direct research to new paths, which is the second goal of medical science. Refusing to conduct the subgroup analysis of large trials would lead to an inefficient use of data, the collection of which has required a substantial amount of resources.

Conclusion

The overall effect of vitamin E on pneumonia risk in the ATBC Study implies that there would be no justification for investing further resources into studying the topic because the narrow confidence interval rejects any substantial overall benefits (RR from 0.88 to 1.14). In contrast, our subgroup analysis suggests a path that should be explored: does vitamin E affect the incidence of pneumonia in physically active males who are nonsmokers or who have had only little exposure to smoking?

Acknowledgments

We thank the ATBC Study (The National Institute for Health and Welfare, Finland and National Cancer Institute, USA) for access to the data.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet*. 2000; 355:1064–1069.
2. Altman DG. Within trial variation–a false trail? *J Clin Epidemiol*. 1998;51:301–303.
3. Freemantle N. Interpreting the results of secondary end points and subgroup analyses in clinical trials: should we lock the crazy aunt in the attic? *BMJ*. 2001;322:989–991.
4. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br J Cancer*. 1977;35:1–39.
5. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol*. 2004;57:229–236.
6. Hemilä H, Virtamo J, Albanes D, Kaprio J. The effect of vitamin E on common cold incidence is modified by age, smoking, and residential neighborhood. *J Am Coll Nutr*. 2006;25:332–339.
7. The ATBC Cancer Prevention Study Group. The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. *Ann Epidemiol*. 1994;4:1–10.
8. 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 cancers in male smokers. *N Engl J Med*. 1994;330: 1029–1035.
9. Hemilä H, Virtamo J, Albanes D, Kaprio J. Vitamin E and beta-carotene supplementation and hospital-treated pneumonia incidence in male smokers. *Chest*. 2004;125:557–565.
10. Powers SK, Jackson MJ. Exercise-induced oxidative stress: cellular mechanisms and impact on muscle force production. *Physiol Rev*. 2008; 88:1243–1276.
11. Hemilä H, Kaprio J, Albanes D, Virtamo J. Physical activity and the risk of pneumonia in male smokers administered vitamin E and β-carotene. *Int J Sports Med*. 2006;27:336–341.
12. Pietinen P, Hartman AM, Haapa E, et al. Reproducibility and validity of dietary assessment instruments: a self-administered food use questionnaire with a portion size picture booklet. *Am J Epidemiol*. 1988;128: 655–666.
13. Saltin B, Grimby G. Physiological analysis of middle-aged and old former athletes. *Circulation*. 1968;38:1104–1115.
14. Hemilä H, Kaprio J. Vitamin E supplementation and pneumonia risk in males who initiated smoking at an early age: effect modification by body weight and vitamin C. *Nutr J*. 2008;7:33.
15. Hemilä H, Kaprio J. Vitamin E supplementation may transiently increase tuberculosis risk in males who smoke heavily and have high dietary vitamin C intake [Discussion: 2009;101:145–147]. *Br J Nutr*. 2008;100:896–902.
16. Hemilä H, Kaprio J. Modification of the effect of vitamin E supplementation on the mortality of male smokers by age and dietary vitamin C. *Am J Epidemiol*. 2009;169:946–953.
17. Hemilä H. Vitamin E is likely to affect mortality even at low doses. *Clin Trials*. 2009;6:392–393.
18. Feinstein AR. The problem of cogent subgroups: a clinicostatistical tragedy. *J Clin Epidemiol*. 1998;51:297–299.

19. Lagakos SW. The challenge of subgroup analyses–reporting without distorting. *N Engl J Med*. 2006;354:1667–1669.
20. Rothwell PM. Treating individuals. Subgroup analysis in randomized controlled trials: importance, indications, and interpretation. *Lancet*. 2005;365:176–186.
21. Bruno RS, Ramakrishnan R, Montine TJ, Bray TM, Traber MG. α-Tocopherol disappearance is faster in cigarette smokers and is inversely related to their ascorbic acid status. *Am J Clin Nutr*. 2005;81: 95–103.
22. Meydani SN, Leka LS, Fine BC, et al. Vitamin E and respiratory tract infections in elderly nursing home residents: a randomized controlled trial [Discussion: 2004;292:2834]. *JAMA*. 2004;292:828–836.
23. Graat JM, Schouten EG, Kok FJ. Effect of daily vitamin E and multivitamin-mineral supplementation on acute respiratory tract infections in elderly persons. *JAMA*. 2002;288:715–721.
24. Hathcock JN, Azzi A, Blumberg J, et al. Vitamins E and C are safe across a broad range of intakes [Discussion: 2005;82:1141–1143]. *Am J Clin Nutr*. 2005;81:736–745.
25. Vandenbroucke JP. Observational research, randomized trials, and two views of medical science. *PLoS Med*. 2008;5:e67.