Title
Statistical design and monitoring of the Carotene and Retinol Efficacy Trial (CARET).

Journal
Controlled clinical trials, 14(4)

ISSN
0197-2456

Authors
Thornquist, MD
Omenn, GS
Goodman, GE
et al.

Publication Date
1993-08-01

DOI
10.1016/0197-2456(93)90228-6

Statistical Design and Monitoring of the Carotene and Retinol Efficacy Trial (CARET)

Mark D. Thornquist, PhD, Gilbert S. Omenn, MD, PhD, Gary E. Goodman, MD, MS, James E. Grizzle, PhD, Linda Rosenstock, MD, MPH, Scott Barnhart, MD, MPH, Garnet L. Anderson, PhD, Samuel Hammar, MD, John Balmes, MD, Martin Cherniack, MD, James Cone, MD, Mark R. Cullen, MD, Andrew Glass, MD, James P. Keogh, MD, Frank Meyskens, Jr., MD, Barbara Valanis, DrPH, and James H. Williams, Jr., MD

Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington (M.D.T.; G.S.O.; G.E.G.; J.E.G.; G.L.A.); University of Washington, Department of Biostatistics, Seattle, Washington (M.D.T.; J.E.G.; G.L.A.); University of Washington, Departments of Environmental Health and Medicine, Seattle, Washington (G.S.O.; G.E.G.; L.R.; S.B.); Swedish Hospital Tumor Institute, Seattle, Washington (G.E.G.); Diagnostic Specialties Laboratory, Bremerton, Washington (S.H.); University of California-San Francisco, Department of Medicine, San Francisco, California (J.B.; J.C.); Yale University, Department of Medicine, New Haven, Connecticut (M.C.; M.R.C.); Kaiser/Portland Center for Health Research (A.G.; B.V.); University of Maryland, Department of Medicine, Baltimore, Maryland (J.P.K.); University of California-Irvine, Department of Medicine and Cancer Center, Orange, California (F.M.,Jr.; J.H.W.,Jr.)

ABSTRACT CARET is a chemoprevention trial of β-carotene and vitamin A with lung cancer as the primary outcome. Participants at high risk for lung cancer are drawn from two populations: asbestos-exposed workers and heavy smokers. The intervention is a daily combination of 30 mg β-carotene and 25,000 IU vitamin A as retinyl palmitate. Nearly 18,000 participants will be followed for a mean 6 years, yielding over 100,000 person-years of follow-up. We project that this sample size will have 80% power to detect a 23% decrease in the incidence of lung cancer cases.

The purpose of this paper is to present the values of the key sample size parameters of CARET; our schemes for monitoring CARET for sample size adequacy, incidence

Address reprint requests to: Mark D. Thornquist, PhD, CARET Coordinating Center, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1124 Columbia Street #MP-859, Seattle, Washington 98104-2092.

Received February 3, 1992; Revised March 30, 1993.

This work is supported by Grant CA 34847 from the National Cancer Institute.

0197-2456/93/$6.00

Controlled Clinical Trials 14:308–324 (1993)

655 Avenue of the Americas, New York, New York 10010

of side effects, and efficacy of the study vitamins; an overview of the data collected; and plans for the primary, secondary, and ancillary analyses to be performed at the end of the trial. These approaches to the design, monitoring, and analysis of CARET are applicable for many other prevention trials.

KEY WORDS: *chemoprevention, lung cancer, sample size, monitoring rules*

INTRODUCTION

The Carotene and Retinol Efficacy Trial (CARET) is a randomized, double-masked, placebo-controlled investigation of the safety and efficacy of β-carotene and vitamin A in decreasing the incidence of lung cancer in two populations at high risk: asbestos-exposed workers and heavy smokers [1,2]. Eligible asbestos-exposed participants are men who are aged 45–69, are current cigarette smokers or have quit within the previous 15 years, have occupational exposure to asbestos beginning at least 15 years previously, and have had significant occupational exposure to asbestos or have chest X rays positive by International Labour Organization criteria for changes consistent with asbestos-related disease. Recruitment of asbestos-exposed participants was restricted to men only because there were very few women in trades with high asbestos exposure more than 15 years ago. Eligible heavy smokers are both men and women who are aged 50–69, are current smokers or have quit within the previous 6 years, and have 20 or more pack-years of cigarette-smoking history. Participants are seen at six study centers: Baltimore, Maryland; Irvine, California; New Haven, Connecticut; Portland, Oregon; San Francisco, California; and Seattle, Washington. Prior to randomization, participants complete a 3-month run-in period receiving placebo medication to determine their adherence to taking the study vitamins and their baseline levels of symptoms, signs, and liver function tests. Participants who take at least 50% of their study vitamins (as determined from weights of returned capsule bottles) and are still eligible are randomized to a daily combination of 30 mg (50,000 IU) β-carotene and 25,000 IU retinol (as retinyl palmitate) dispensed in a single capsule or to a comparable capsule containing placebo medication. Randomization is by household, stratified by population (asbestos-exposed participants and heavy smokers) and study center, using blocked randomization with random block size to ensure balance of follow-up time in the two arms. Households are randomized because of the possibility of participants taking capsules from their spouses' bottles.

In the first year after randomization, participants are contacted twice by telephone (at 3 and 9 months after randomization) and visit twice at the study center (at 6 and 12 months). Subsequently, participants are contacted twice by telephone and once at a study center visit each year. At each contact there is monitoring for possible toxicity, reflected in symptoms, signs, or laboratory values potentially attributable to the study vitamins.

CARET builds on the experience of two pilot studies performed in Seattle (1985–1988) that demonstrated that the recruitment strategies could accrue targeted numbers of participants, that adherence both to taking the study vitamins and to the data collection schedule was excellent, and that side effects were few. The first pilot study initiated a phase III trial of the safety and efficacy of the study vitamins in 816 asbestos-exposed participants random-

ized to a daily combination of 15 mg β-carotene and 25,000 IU retinol or a placebo medication [3]. Participants were eligible up to age 74 and were not required to have a history of cigarette smoking; otherwise, the eligibility criteria were the same as for the asbestos-exposed population in CARET [1]. The second pilot study was a phase II trial of the comparative safety of the study vitamins in heavy smokers [4]. The eligibility criteria were identical to those for heavy smokers in CARET. We randomized 539 men and 490 women to one of four intervention groups: a daily combination of 30 mg β-carotene and 25,000 IU retinol; 30 mg β-carotene only; 25,000 IU retinol only; or placebo medication. All 1845 participants in the two pilot studies continue to be followed for outcomes in CARET, together with approximately 16,000 additional participants, called "efficacy" participants, whom we are currently recruiting. With the expansion to the full-scale trial, all participants in the pilot studies who were receiving any active vitamins had their dosages changed, where necessary, to match the dosage used in CARET. Only those originally randomized to placebo medication continue to receive placebo medication. As a result, three fourths of the participants from the pilot study in heavy smokers are in the group randomized to receive active vitamins. Participants from the pilot studies are monitored more closely for indications of toxicity than are the efficacy participants, since due to their longer follow-up time toxicities related to cumulative dose should appear in them years before they would appear in the efficacy participants.

The primary trial outcome is lung cancer. Secondary outcomes are malignant mesothelioma (in the asbestos-exposed population), other malignancies, and death. Since we do not anticipate that vitamin A and β-carotene will be effective against existing lung tumors, we down-weight lung cancers diagnosed in the first 2 years after randomization. The primary analysis of the effectiveness of the study vitamins will involve a comparison between the intervention groups of the numbers of weighted lung cancers, with weight function $w(t) = \min(t/2,1)$, where t is the time in years since randomization at which the lung cancer is diagnosed. To achieve adequate power to test the effect of vitamin A and β-carotene on the number of weighted lung cancers, CARET has a goal of randomizing approximately 17,700 participants.

The purpose of this paper is to present the key parameters of the statistical design of CARET, as a guide to researchers designing other chemoprevention trials. We detail the determinants of the number of participants needed and the duration of the trial, describe our four schemes for monitoring progress, and outline the data collected and the analyses planned.

ESTIMATION OF SAMPLE SIZE AND TRIAL DURATION

The estimation of sample size requires the specification of the incidence of the trial outcome (lung cancer), adherence to taking the study vitamins, and rate of loss to follow-up, together with statistical parameters. Table 1 presents the values of the sample size parameters utilized in the design of CARET and reported by Omenn et al. [1]. These values reflect our review of the relevant literature and our pilot study experiences [3,4].

We developed models for the incidence rate of lung cancer as a function of risk factors (age, gender, smoking history, and asbestos exposure) in the

Table 1 Parameters Used in the Sample Size Calculations for CARET

	Initial Values[a]			Updated Values[b]		
Lung cancer rate (per 10^5)	0 yr	8 yr		0 yr	8 yr	
Pilot asbestos	600	1100		300	500	
Efficacy asbestos	800	1300		800	1000	
Smokers	300	600		400	600	
Capsule consumption rate, active vitamin group	0 yr	3 yr	8 yr	0 yr	3 yr	8 yr
Pilot asbestos	100%	80%	75%	98%	73%	64%
Pilot smokers	100%	67%	62%	98%	67%	58%
Efficacy asbestos	100%	80%	75%	94%	85%	70%
Efficacy smokers	100%	67%	62%	95%	79%	54%
Maximum supplementary vitamin intake, placebo medication group	0 yr 0%	3 yr 5%	8 yr 5%	0 yr 0%	3 yr 5%	8 yr 5%
MPCE[c]	33%			33%		
Time lag to full effectiveness	2 years			2 years		
Loss to competing risks						
Death rate (per 10^5)	0 yr	8 yr		0 yr	8 yr	
Pilot asbestos	1900	2600		1000	2000	
Efficacy asbestos	1900	2600		900	1700	
Smokers	1500	2500		700	1600	
Loss to follow-up	2% over trial			2% over trial		
Duration of accrual period						
Pilots	3 years			3 years		
Efficacy	5 years			Table 3		
Type I error	.05 two-sided			.05 two-sided		
Power	80%			80%		
Number of participants	4,277 Asbestos 13,629 Smokers 17,906 Total			3,983 Asbestos 13,719 Smokers 17,702 Total		

[a]Reproduced from Omenn et al. [1].

[b]Based on monitoring of the duration required for desired power through June 1992 (see text).

[c]Maximal potential chemopreventive effect, the reduction in risk in fully adherent participants on the active vitamins compared to placebo.

absence of an intervention using 1981–1985 lung cancer incidence rates by age and gender from the Surveillance, Epidemiology, and End Results Program [5]; relative risks for current smokers to never-smokers and to former smokers by gender from the American Cancer Society's CPS-II study [6]; and distributions of lung cancer risk factors by age and gender in the U.S. from the 1985 National Health Interview Survey [7] and the 1986 Adult Use of Tobacco Survey [8]. We used a relative risk of 3 for the effect of asbestos-exposure, based on other studies in similar populations [9–11]. The lung cancer incidence rates expected in the CARET participants were projected from these models using the distributions of risk factors in the randomized participants. Lung cancer and death rates increase by time since randomization due to the aging of the study participants.

Adherence to taking the study vitamins is operationally measured by the

mean number of capsules of study vitamins taken per day by participants (called the *capsule consumption rate*). The mean is over all participants who are still alive, including those who have stopped taking study vitamins for any reason, and is based on weights of returned bottles of study vitamins and on self-reports when bottle weights are not available. Capsule consumption rates for the first 3 years after randomization (Table 1) were based on experience from the pilot studies. These rates were similar to those reported for three large cardiovascular trials with cholesterol-lowering drugs [12], aspirin [13], and β blockers [14]. We projected a decline in the capsule consumption rate of 1% per year for years 3–8, based on the decline seen in later years in the Lipid Research Clinics trial [12].

The CARET design allows for participants in the group receiving placebo medication to be taking up to an equivalent of 5% of the specified dose through intake of supplementary vitamins. We expect this factor to be conservative; through June 1992 less than 1% of participants were taking supplemental vitamin A over 5,500 IU/day or any β-carotene.

We planned CARET with the view of being able to examine the potential efficacy of vitamin A and β-carotene in two high-risk populations: asbestos-exposed workers and heavy smokers. Our target was to have 80% power to detect a 33% maximal potential chemopreventive effect (MPCE) in the total study population and a 50% MPCE in each subpopulation (asbestos-exposed participants and heavy smokers) separately. The MPCE is defined to be the reduction in lung cancer risk in the intervention arm if all participants in both arms were fully adherent. Because of the time lag to full effectiveness and the allowance for less than full adherence of the trial participants, the actual observable reductions will be less than the MPCE.

We developed models for the incidence rate of death in the absence of an intervention as a function of relevant risk factors (age, gender, smoking history, and asbestos exposure) from 1980 U.S. total mortality by age and gender [16]; relative risks for current smokers to never-smokers and to former smokers by gender from the U.S. Surgeon General's report [17]; and the distributions of lung cancer risk factors in the United States used above in the modeling of lung cancer incidence rates. Total mortality among CARET participants was projected, and then the lung cancer incidence rate determined above was subtracted to estimate the loss to causes of death other than lung cancer. Lung cancer incidence is a good approximation to lung cancer mortality because of low survival after the diagnosis of lung cancer—less than 15% after 5 years [5].

We established study centers in five regions with major cohorts of asbestos-exposed shipyard workers (Baltimore, New Haven area, Portland, and San Francisco, in addition to the original study center in Seattle) and, based on pilot experience of response to recruitment efforts, set requirement targets of 4277 randomized asbestos-exposed participants (including the participants from the pilot study in asbestos-exposed workers). We targeted the accrual of 13,629 heavy smokers (including the participants from the pilot study in heavy smokers) at study centers in Irvine, Portland, and Seattle. For this sample size, given the parameters outlined in Table 1, the duration of follow-up needed to achieve our desired power was 7.8 years from the first ran-

domizations to the efficacy cohort in June 1989. Allowing an additional 6 months of follow-up for reporting lags, follow-up would conclude in September 1997, if all these projections were met.

Although the sample size estimate depends on a large number of factors (Table 1), our estimates have been robust across the alternatives that we have explored. The product of the sample size and duration of the trial, approximately 100,000 person-years of follow-up, depends primarily on the incidence rate of lung cancer in the group receiving placebo medication and on the anticipated reduction in the incidence rate of lung cancer resulting from our intervention. The incidence rate in the group receiving placebo medication depends on the distribution of risk factors in that group. The anticipated reduction in the incidence rate is influenced by the amount of study vitamins received (measured by the capsule consumption rate) and the time lag to full effectiveness. The number of person-years is insensitive to changes in the accrual schedule or loss to competing risks. The other factors in Table 1 are considered fixed for CARET.

Smoking cessation has been estimated to have only a small effect on sample size. Under the incidence rate models developed for the CARET sample size estimation, even if 10% of currently smoking participants stopped smoking each year with no relapse, the duration of the trial would change by only 1 month. Life table estimates of the net reduction (quitters minus relapsers) in the number of efficacy participants who are current smokers, based on self-reports of smoking behavior through June 1992, are about 6% per year in both populations. Thus, the current rate of reduction in smoking observed in CARET has negligible effect on the expected duration of the trial. Assistance with smoking cessation is a standardized procedure in CARET.

MONITORING

Large prevention trials such as CARET must be monitored closely to justify the investment of the sponsoring agency. It is important to demonstrate that the trial is progressing as planned, or, if it is not, to identify how to make timely changes. In addition to the conduct of cost analyses described in the accompanying paper [18], we monitor the progress of CARET in four ways (Table 2). The distribution of risk factors, adherence to taking the study vitamins, and the overall incidence rate of lung cancer are monitored to determine whether the initial design assumptions are being met; if not, the duration of the trial is adjusted to ensure that the trial maintains the desired power. Symptoms, signs, and abnormal laboratory results potentially attributable to β-carotene and retinol are monitored to determine whether our dosage of the study vitamins results in toxicity, which if greater than any potential benefit of the study vitamins in preventing cancer could result in termination of the trial with the conclusion of no net benefit. The difference between the groups receiving active and placebo medication in the incidence rates of lung cancer is monitored so that the trial may be stopped early if there is sufficient evidence of efficacy of the study vitamins. Finally, if requested by the Safety and Endpoints Monitoring Committee (SEMC) of CARET, the conditional power of the trial will be monitored; if low, the trial may be terminated with the

Table 2 CARET Monitoring Schemes

Factor Monitored	Reason	Frequency	Scheme
Lung cancer risk factors in randomized population	Ensure power is maintained	Twice per year	Input updated values into sample size formulas
Incidence of potential side effects	Ensure safety of participants, quantify side effects due to vitamins, stop study if serious problem of side effects	Twice per year	Review by external Safety and Endpoints Monitoring Committee (SEMC)
Lung cancer incidence	Stop study early if strong evidence of efficacy	Interim analyses at 1/3 and 2/3 of weighted[a] endpoints	O'Brien–Fleming boundaries [25]
Conditional power	Stop study if low power for demonstrating efficacy	As requested by SEMC	Stochastic curtailment [26]

[a]Weighted by $w(s) = \min(s/y, 1)$, where s is the time since randomization and y is the time lag to full effectiveness (2 years).

conclusion of no evidence of efficacy of the study vitamins. These four monitoring schemes are described below.

Monitoring Duration Required for Desired Power

The duration of CARET depends on the rate of recruitment, the incidence rates of lung cancer and death in the study population, and the amount of study vitamins taken by participants. The sample size calculations outlined in the previous section used estimates of these factors based on our pilot study experience and experience from other prevention trials. Every 6 months, we update the sample size calculation using baseline information from the growing number of randomized participants: actual accrual schedule, mean age, mean number of cigarettes smoked per day, and mean number of years since quitting smoking for former smokers. We also update capsule consumption rates estimated from weights of returned vitamin bottles and self-reports, and examine the rates of lung cancers and deaths in the study population. We do not analyze the reduction in the incidence rate of lung cancer resulting from our intervention in order to avoid the need for adjustment for multiple tests. The estimated duration of the trial is recomputed, and the results are reviewed by the CARET principal investigators for any effect on trial plans, including budgets.

As an example of the results of this monitoring in CARET, the right column of Table 1 presents the trial design parameters updated through June 1992. The lung cancer incidence rates in asbestos-exposed participants, capsule consumption rates, death rates, and numbers of participants all were revised from the original parameters. We have had to increase the planned duration of CARET by 0.5 year in order to maintain our power at 80% to detect the

alternative hypothesis. The effects of each individual change are described below.

The lung cancer incidence rate in asbestos-exposed participants was lower than projected, at least in part due to the removal of prevalent cases by the chest X-ray screening for asbestos-related changes performed prior to randomization (14 cases confirmed and 7 possible). As of June 1992 the observed lung cancer incidence rates were 45% of that projected for efficacy asbestos-exposed participants and 75% of that projected for the participants from the pilot study in asbestos-exposed workers, who had had longer follow-up. For the updated values in Table 1, we adjusted the lung cancer rates for the asbestos-exposed participants, so that by the end of the trial the observed rate would be 80% of the originally projected rates. We expect this projection to be conservative. This change increased the projected duration of the trial by 0.6 year.

The capsule consumption rates were updated using observed rates for initial years and projecting future years assuming a linear decline using the slope observed in the last year for which we have data (rather than the 1%-per-year decline for years 3–8 originally assumed). We expect that the rate of decline will decrease in future years of the trial, so the assumption of linear decline is conservative. The capsule consumption rates at the time of randomization reflect the actual rates observed during the 3-month enrollment period with placebo medication prior to randomization. The changes in capsule consumption rates decreased the projected trial duration by 0.2 year; the observed rates in the first years after randomization have been higher than originally projected, which offset the eventually lower rates resulting from the linear projection.

Through June 1992, rates of death in the study population were 60% of projected; the study population actually had a lower death rate than the entire U.S. population for the trial's age range, even though all trial participants had asbestos exposure and/or a history of cigarette smoking. A likely explanation is a healthy volunteer effect. A reduction of the same magnitude, with several more years of follow-up, has been observed in the Alpha-Tocopherol Beta-Carotene chemoprevention trial in smokers being conducted in Finland (personal communication). Thus, we expect that the lower rates of death may persist for the duration of the trial. In Table 1, the projected death rates have been multiplied by 0.6; this decreased the projected duration of the trial by 0.2 year.

Changes through June 1992 in the accrual schedules for asbestos-exposed participants included a slower rate of accrual than originally planned and a reduced target for one study center because of poor response to recruitment efforts, but the other four study centers exceeded their accrual targets. Changes in the accrual schedules for heavy smokers include a slower rate of accrual than originally planned and the addition of recruitment of smoker-eligible spouses at the study centers that originally randomized only asbestos-exposed participants. These changes increased the projected duration of the trial by 0.3 year. The accrual for each CARET study center (actual through June 1992 and targets thereafter) is presented in Table 3.

Based on the parameters revised in June 1992, the projected duration of CARET is 8.3 years from the first randomization; with 6 additional months

Table 3 CARET Accrual Schedules

	Year of Randomization										
Study Center	85–86[a]	86–87	87–88	88–89	89–90	90–91	91–92	92–93	93–94	94–95	Total
Asbestos-Exposed Workers											
Baltimore	—	—	—	—	187	454	165	—	—	—	806
New Haven	—	—	—	—	267	385	238	110	—	—	1,000
Portland	—	—	—	—	23	98	100	69	—	—	290
San Francisco	—	—	—	—	206	291	253	120	—	—	870
Seattle pilot	190	444	182	—	—	—	—	—	—	—	816
Seattle efficacy	—	—	—	—	14	23	119	45	—	—	201
Total[b]	190	444	182	—	697	1,251	875	344	—	—	3,983
Heavy Smokers											
Baltimore	—	—	—	—	—	—	—	30[c]	—	—	30
Irvine	—	—	—	—	—	—	328	1,820	1,986	166	4,300
New Haven	—	—	—	—	—	—	—	30[c]	—	—	30
Portland	—	—	—	67	778	1,295	998	862	—	—	4,000
San Francisco	—	—	—	—	—	—	—	30[c]	—	—	30
Seattle pilot	378	499	152	—	—	—	—	—	—	—	1,029
Seattle efficacy	—	—	—	50	1,205	690	1,456	899	—	—	4,300
Total[b]	378	499	152	117	1,983	1,985	2,782	3,671	1,986	166	13,719

[a]Years run from July to June to match funding schedule.

[b]Data from 85–86 through 91–92 are actual accruals of randomized participants; later periods are target numbers.

[c]Additional recruitment decided on in 1992 as a retention initiative, inviting smoker-eligible spouses in the asbestos-only study centers.

Table 4 Projected Person-years of Follow-up and Numbers of Lung Cancers for CARET

	Active Vitamin Group			Placebo Medication Group		
Population and Cohort	Person-years of Follow-up	Numbers of Lung Cancers	Numbers of Weighted[a] Lung Cancers	Person-years of Follow-up	Numbers of Lung Cancers	Numbers of Weighted Lung Cancers
Pilot asbestos	4,000	16	15	4,000	20	18
Pilot smokers[b]	7,800	31	28	2,600	13	12
Efficacy asbestos	9,700	65	54	9,700	85	73
Efficacy smokers	34,400	113	92	34,300	144	121
Total	55,900	225	189	50,600	262	224

[a]Weighted by $w(s) = \min(s/y, 1)$ where s is the time since randomization and y is the time lag to full effectiveness (2 years).

[b]Pilot smokers are randomized 3/4 to active, 1/4 to placebo capsules; all other subgroups are randomized 1/2 to active, 1/2 to placebo.

for reporting and confirming of lung cancers, follow-up is planned to conclude in April 1998. Table 4 presents the projected person-years of follow-up and numbers of lung cancers expected. Table 5 presents the anticipated reductions in lung cancer incidence, ranging from 21% in the pilot heavy smokers to 25% in the pilot asbestos-exposed participants. The observable chemopreventive efficacy is reduced from the 33% MPCE (Table 1) by our assumptions

Table 5 Projected Chemopreventive Effect to be Achieved in CARET

Population and Cohort	All Lung Cancers (%)	Weighted Lung Cancers (%)
Pilot asbestos	21	22
Pilot smokers	19	20
Efficacy asbestos	23	26
Efficacy smokers	22	24
Total	22	23

about the capsule consumption rates and the time lag to full effectiveness. For the two high-risk populations, CARET is projected to be able to detect a 50% MPCE with 85% power in the asbestos-exposed participants alone and with 89% power in the heavy smokers alone.

The power of this trial to detect reductions in the rate of death from all causes likewise depends on the assumed time lag to full effectiveness. CARET would have 80% power to detect an MPCE for the death outcome of 17% if the effect of the study vitamins on the risk of death were immediate; 19% if the time lag is 2 years; and 21% if the time lag is 4 years. To detect a reduction in the rate of mesothelioma with a power of 80%, we estimate that the MPCE would have to be 58%, based on an estimated incidence rate of 6 per 1000 person-years and a time lag of 2 years.

Other outcomes of interest include prostate cancer and atherosclerotic disease. Virtamo and Huttunen [19] concluded that the data on the association between dietary β-carotene or vitamin A and prostate cancer are equivocal, with some studies suggesting a protective effect and some an adverse effect. CARET has an 80% power to detect a 36% reduction or a 44% increase in the incidence of prostate cancer, assuming a time lag of 2 years. Ridker et al. [20], in an abstract of preliminary findings from the Physicians Health Study, advocated a prospective trial of β-carotene for atherosclerotic outcomes; CARET has 80% power to detect a 19% MPCE on the incidence of myocardial infarction plus sudden cardiac death, assuming no time lag.

Monitoring for Side Effects

Large doses of retinol may lead to hypervitaminosis A, resulting in skin changes, psychological affective disorders, bone pain, and liver damage [21]. The safety of chronic doses of 25,000 IU/day must be investigated, to extend the experience from the pilot studies [3,4]. There are case reports of hypervitaminosis A in adults at doses as low as 25,000 IU per day 7–10 years [22,23]. The only known side effect of β-carotene is yellowing of the skin [24].

The CARET protocol includes monitoring all participants for 13 symptoms or signs that may be indicative of vitamin A toxicity at each contact [1], and monitoring serum levels of glutamic-oxaloacetic transaminase (GOT), alkaline phosphatase, total cholesterol, and triglycerides annually in participants continuing from the pilot studies. Symptoms that exceed predefined levels of severity are managed with a scheme that includes stopping the study vitamins and rechallenging [4].

Every 6 months the CARET Coordinating Center analyzes the accumulated data on symptoms, signs, and laboratory values and reports the results by coded intervention group to the SEMC. Analyses of data from active participants compare mean symptom grades and percentages above threshold grades; analyses of data from participants who have stopped taking the study vitamins (called "inactive" participants) examine reasons for and time course to becoming inactive, and means and maximum symptom grades immediately prior to becoming inactive. Through June 1992, differences between the intervention groups were negligible. In the event the differences become significant, the SEMC may recommend additional analyses and/or changes in the trial. We have not set up formal monitoring rules for side effects because of the large number monitored. A rule based on a combined score across all symptoms and signs monitored would not be sensitive to large differences in any one symptom or sign, while separate rules for each symptom and sign would result in extreme stopping boundaries because of multiple testing. The SEMC prefers to synthesize the information on symptoms and signs with evidence of efficacy in deciding by formal vote whether to continue the trial at each semiannual meeting.

Monitoring of Outcomes

Outcomes in CARET are lung cancer, mesothelioma, other cancers (except nonmelanoma skin cancers), and death. Sources of initial reports of outcomes are trial participants or their next of kin, cancer registries, state boards of health, and the National Death Index. All initial reports of outcomes are reviewed by CARET's internal endpoints committee to determine whether they are indeed study outcomes. Slides and specimens from all tumors with lung involvement are reviewed by CARET's consulting pathologist (S.H.), a specialist in pulmonary immunopathology and mesotheliomas. For each cancer, the committee comes to a decision on the primary site, histology, date of histological diagnosis, and relationship of the cancer to death if the participant has died. For each death, the committee decides on the cause (immediate and underlying), date of death, and whether cancer was present at death. Details on all outcomes, both confirmed and pending confirmation, are reported semiannually to the SEMC.

The formal CARET monitoring for stopping the trial due to demonstrated efficacy of the study vitamins is provided by O'Brien–Fleming boundaries [25]. Interim analyses will be performed at one third and two thirds of the projected number of weighted lung cancers. Using O'Brien–Fleming boundaries, there is little probability of stopping early unless there is strong evidence of efficacy of the study vitamins, since we also need to establish the safety of long-term supplementation at the trial dosage. The critical p values are .0006 for the first interim analysis, .015 for the second, and .047 for the final analysis [25]. Based on experience through June 1992, we project that one third of the weighted outcomes will occur by January 1994 and two thirds by January 1996. Allowing 6 months for reporting and confirming lung cancer outcome cases, we expect to perform the scheduled interim analyses in the summers of 1994 and 1996.

Monitoring of Conditional Power

In addition to stopping the trial for toxicity or for efficacy, we allow for stopping the trial because of the absence of a difference between the intervention groups in the incidence of lung cancer. We will use the method of stochastic curtailment [26], in which the power for detecting the desired effect at the end of the trial is computed under the alternative hypothesis conditional on the data to date. Such computations will be performed at the request of the SEMC. We have no formal stopping boundaries because the decision on whether to continue the trial in the event of low projected power would depend on other factors, such as the incidence and severity of side effects, cost, and results of other trials.

PLANS FOR ANALYSIS

Data Collected

The schedule of CARET participant contacts is described in the Introduction. The following data are collected:

- *First Visit:* Basic demographics (age, gender, race, education, mobility), smoking history, occupational history, family cancer history, general health, health history, symptoms and signs, serum analytes, food frequency questionnaire, reasons for participation, and (for asbestos-exposed participants only) American Thoracic Society (ATS) questionnaire for respiratory symptoms [27], spirometry [28], and chest X ray.
- *Follow-up visits:* Vital signs (height, weight, blood pressure), symptoms and signs, supplemental vitamin use, study vitamin adherence, reasons for not taking study vitamins (if participant has missed at least 10 capsules), general health, worry about health, updated cigarette-smoking history, actions taken by trial personnel to assist the participant to quit smoking, reasons for participating in the trial, effects on the participant of participation, and perceptions of clinical improvements possibly related to the study vitamins.
- *Follow-up telephone calls:* Symptoms and signs, supplemental vitamin use, reasons for not taking study vitamins, and new medical conditions.
- *Every 2 years:* Food frequency, serum (annually for pilot participants), and (for asbestos-exposed participants only) ATS questionnaire and spirometry.

Inactive participants, who are no longer taking study vitamins, are contacted by telephone twice every year, at which times vital status, occurrence of any cancer, current smoking status, average number of cigarettes smoked per day, and general health are assessed.

Blood Collection and Serum Analysis

Sufficient blood to yield four 1.7-ml aliquots of serum is collected from pilot participants every year and from efficacy participants at their first visit and every 2 years after randomization. We collect an additional two 1.7-ml

aliquots of serum from a 12.5% cohort of participants for the CARET serum laboratory quality assurance. Serum aliquots are stored at −20°C at the study centers for up to 2 weeks prior to shipment in dry ice to the Coordinating Center, where they are then stored at −70°C. The laboratory quality assurance includes participation in the National Institute of Standards and Technology round-robins and analysis of serum from standard pools with each run. In addition, 1 of every 10 samples sent to the laboratory for analysis is a masked split duplicate of another sample in that batch. The Coordinating Center monitors the performance on the split duplicates to assess within- and between-run variability. Blood spots on Schleicher and Schuell no. 903 filter papers [29,30] will be obtained once in 1993 to save for DNA analyses by polymerase chain reaction methods as appropriate probes become available.

For economy we have organized our analysis for the serum around the case-cohort method developed by Prentice [31]. The cohort analyzed for comparison with the cases consists of all of the participants from the pilot studies and 10% of the efficacy participants, masked to the study centers and chosen independently of the 12.5% quality assurance cohort. One aliquot from each sample from the cohort will be analyzed as received for levels of retinol, retinyl palmitate, α-carotene, and β-carotene by a modification of the method of Milne and Botnen [32]. A single analysis typically uses 0.5–0.7 ml of serum; the remaining serum is refrozen at −70°C. Stored serum samples are also analyzed for those participants having trial outcomes. We will have approximately 7 controls per case for the lung cancer outcomes and 2.5 controls per case for the death outcomes. As a result, the case-cohort method has 90% of the efficiency for the detection of the effects of serum levels on lung cancer risk (75% for the effects on mortality risk) at only 20% of the cost, compared to analyzing serum on all participants.

Data Analysis

There are two specific aims of CARET: to test the efficacy of the study vitamins in preventing lung cancer and to test the safety of our dosage of the study vitamins. The test of efficacy of the study vitamins is performed by a weighted log-rank statistic, stratified on study center, population (asbestos-exposed participants and heavy smokers), and cohort (pilots and efficacy). Lung cancers diagnosed in the first 2 years after randomization are down-weighted linearly to account for an expected time lag between starting to take the active vitamins and observation of the full chemopreventive effect due to the possibility of undiagnosed lung tumors in some participants at randomization. Sample size for the weighted log-rank statistic is based on the method of Self et al. [33], modified to account for the weighting in the computation of the variance of the test statistic (details available on request from the authors).

The test of safety is performed by a comparison between the intervention groups in the trends and severity of relevant symptoms, signs, and liver function tests, using the method of Laird and Ware [34]. This test has greatest power for the specific alternative of a difference between arms in the slopes of the data, but should have adequate power to detect effects that plateau.

Secondary analyses will examine the relationship between the study vitamins and the other outcomes (mesothelioma, other malignancies, and

death), and the adjustment of the primary and secondary analyses for potential covariates, including:

- Lung cancer risk factors: age, gender, cigarette smoking history and cessation success, serum analyte levels, and asbestos exposure
- For the asbestos-exposed population, baseline and follow-up measures of lung disease: X ray, spirometry, ATS questionnaire

Secondary analyses will also be performed to better quantify the time pattern of any symptoms, signs, or abnormal laboratory measures found to be associated with the intervention. The relationship between adherence and outcome measures is of interest; however, recent work by Efron and Feldman [35] points out that the parameters of greatest interest are not identifiable, and so can be estimated only by making untestable assumptions.

Ancillary analyses will address the effect of the intervention on other trial measures and the associations among the measures. An example of the former is a test of the effect of the intervention in asbestos-exposed participants on lung function loss as assessed by the serial spirometries and the ATS questionnaire. An example of the latter is the predictive power of baseline X rays on subsequent lung function loss. CARET will also have good power to analyze effects of the intervention on cardiovascular outcomes, as noted above.

CONCLUSION

CARET is a major cancer chemoprevention trial testing the effects of β-carotene and retinyl palmitate on the incidence of lung cancer in individuals at high risk for the disease. We will randomize nearly 18,000 participants and amass over 100,000 person-years of follow-up to achieve our goal of 80% power to detect a maximal potential chemopreventive effect of 33% and an observable reduction of 23% in lung cancer incidence. CARET, the Physicians Health Study [36], and the Alpha-Tocopherol Beta-Carotene Study [37] of 29,000 smokers in Finland are the National Cancer Institute's three major cancer chemoprevention trials utilizing β-carotene plus vitamin A, β-carotene alone, and β-carotene plus vitamin E, respectively. We expect these trials together to clarify the roles of these vitamins as cancer chemopreventive agents.

Preventing cancer is preferable to needing to treat cancer. Epidemiological and animal studies have been valuable in identifying substances or activities that may help in preventing cancer. We must progress beyond the limitations of these studies to rigorous trials of such substances or activities, with painstaking effort in design, recruitment, retention, and monitoring to achieve the designed power. When executed properly, prevention trials represent an essential test of the hypotheses generated by epidemiological studies with their potential for bias. Zelen [15] has argued that primary prevention trials for cancer outcomes are not feasible; we are demonstrating with CARET that they are.

REFERENCES

1. Omenn GS, Goodman G, Grizzle J, Thornquist M, Rosenstock L, Barnhart S, Anderson G, Balmes J, Cherniack M, Cone J, Cullen M, Glass A, Keogh J, Meyskens F, Valanis B: CARET, the beta-carotene and retinol efficacy trial to prevent

lung cancer in asbestos-exposed workers and in smokers. Anti-Cancer Drugs 2:79–86, 1991

2. Omenn GS, Goodman G, Grizzle J, Thornquist M, Rosenstock L, Barnhart S, Anderson G, Balmes J, Cherniack M, Cone J, Cullen M, Glass A, Keogh J, Valanis B, Williams J Jr: Recruitment for CARET, the beta-carotene and retinol efficacy trial to prevent lung cancer in smokers and in asbestos-exposed workers. Western J Med 156:540–544, 1992
3. Omenn GS, Goodman GE, Thornquist MD, Rosenstock L, Barnhart S, Gylys-Colwell I, Metch B, Lund B: CARET, the Carotene and Retinol Efficacy Trial to prevent lung cancer in high risk populations: pilot study with asbestos-exposed workers. In press
4. Goodman GE, Omenn GS, Thornquist MD, Lund B, Metch B, Gylys-Colwell I: CARET, the Carotene and Retinol Efficacy Trial to prevent lung cancer in high risk populations: pilot study with cigarette smokers. In press
5. Sondik EJ, Young JL, Horm JW, Gloeckler Ries, LA: 1986 Annual cancer statistics review. NIH Publication No. 87-2789, U.S. Department of Health and Human Services, 1987
6. Stellman SD, Garfinkel L: Smoking habits and tar levels in a new American Cancer Society prospective study of 1.2 million men and women. JNCI 76:1057–1063, 1986
7. National Cancer for Health Statistics: Provisional data from the health promotion and disease prevention supplement to the National Health Interview Survey: United States, Jan–March 1985. Advance data from vital and health statistics, No. 113. DHHS Publication No. (PHS) 1985:86–1250, 1985
8. U.S. Department of Health and Human Services: Tobacco Use in 1986. Methods and basic tabulations from Adult Use of Tobacco Survey. Public Health Service, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. DHHS Publication No. (OM)90–2004, 1990
9. Nicholson WJ, Perkel G, Selikoff IJ: Occupational exposure to asbestos: population at risk and projected mortality: 1980–2030. Am J Ind Med 3:259–311, 1982
10. Selikoff IJ, Lilis R, Nicholson WJ: Asbestos disease in United States shipyards. Ann NY Acad Sci 330:295–311, 1979
11. Selikoff IJ, Hammond EC, Seidman H: Mortality experience of insulation workers in the United States and Canada. Ann NY Acad Sci 330:91–116, 1979
12. Lipid Research Clinics Program: The lipid research clinics coronary primary prevention trial results. 1. Reduction in incidence of coronary heart disease. JAMA 251:351–364, 1984
13. Aspirin Myocardial Infarction Study Research Group: A randomized, controlled trial of aspirin in persons recovered from myocardial infarction. JAMA 243:661–669, 1980
14. Beta-Blocker Heart Attack Trial Research Group: A randomized trial of propranolol in patients with acute myocardial infarction. JAMA 247:1709–1714, 1982
15. Zelen M: Are primary cancer prevention trials feasible? JNCI 80:1442–1444, 1988
16. National Center for Health Statistics: Vital statistics of the United States, 1980, Vol 2, Mortality, Part B: DHHS Publication No. (PHS) 85-1102. Public Health Service, Washington, DC. U.S. Government Printing Office, 1985
17. U.S. Public Health Service: The health consequences of smoking: a public health service review—1967. U.S. Department of Health, Education, and Welfare, Public Health Service. PHS Publication No. 1696, revised 1968
18. Thornquist MD, Urban N, Tseng A, Edelstein C, Lund B, Omenn GS: Research cost analyses to aid in decision making in the conduct of a large prevention trial, CARET. This issue
19. Virtamo J, Huttunen JK: Beta-carotene, vitamin A and prostate cancer. Ann Med 24:143–144, 1992

20. Ridker PM, Manson JE, Buring JE, Goldhaber SZ, Hennekens CH: Clinical characteristics of nonfatal myocardial infarction among individuals on prophylactic low-dose aspirin therapy. Circulation 84:708–711, 1991
21. Hathcock JN, Hattan DG, Jenkins MY, McDonald JT, Sundaresan PR, Wilkening VL: Evaluation of vitamin A toxicity. Am J Clin Nutr 52:183–202, 1990
22. Weber FL Jr, Mitchell GE Jr, Powell DE, Reiser BJ, Banwell JG: Reversible hepatotoxicity associated with hepatic vitamin A accumulation in a protein-deficient patient. Gastroenterology 82:118–123, 1982
23. Hatoff DE, Gertler SL, Miyai K, Parker BA, Weiss JB: Hypervitaminosis A unmasked by acute viral hepatitis. Gastroenterology 82:124–128, 1982
24. Thomsen K, Schmidt H, Fischer A: Beta-carotene in erythropoietic protoporphyria: 5 years' experience. Dermatologica 159:82–86, 1979
25. O'Brien PC, Fleming TR: A multiple testing procedure for clinical trials. Biometrics 35:549–556, 1979
26. Lan KKG, Simon R, Halperin M: Stochastically curtailed tests for long-term clinical trials. Commun Stat—Sequential Analysis 1:207–219, 1982
27. Comstock GW, Tockman MS, Helsing KJ, Hennesy KM: Standardized respiratory questionnaires: comparison of the old with the new. Am Rev Resp Dis 119:45–53, 1979
28. Eisen E: Standardizing spirometry: Problems and prospects. In Rosenstock L. (ed): Occupational Pulmonary Disease. State of the Art Reviews. Philadelphia, Hanley and Belfus, 1987
29. Cassol S, Salas T, Arella M, Neumann P, Schechter MT, O'Shaughnessy M: Use of dried blood spot specimens in the detection of human immunodeficiency virus type 1 by the polymerase chain reaction. J Clin Microbiol 29:667–671, 1991
30. Jinks DC, Minter M, Tarver DA, Vanderford M, Hejtmancik JF, McCabe ER: Molecular genetic diagnosis of sickle cell disease using dried blood specimens on blotters used for newborn screening. Human Genet 81:363–366, 1989
31. Prentice RL: A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika 73:1–11, 1986
32. Milne DB, Botnen J: Retinol, alpha-tocopherol, lycopene, and alpha- and beta-carotene simultaneously determined in plasma by isocratic liquid chromatography. Clin Chem 32:874–876, 1986
33. Self S, Prentice R, Iverson D, Henderson M, Thompson D, Byar D, Insull W, Gorbach SL, Clifford C, Goldman S, Urban N, Sheppard L, Greenwald P: Statistical design of the Women's Health Trial. Controlled Clin Trials 9:119–136, 1988
34. Laird NM, Ware JH: Random-effects models for longitudinal data. Biometrics 38:963–974, 1982
35. Efron B, Feldman B: Compliance as an explanatory variable in clinical trials. J Am Stat Assoc 86:9–17, 1991
36. Hennekens CH, Eberlein K: A randomized trial of aspirin and beta-carotene among U.S. physicians. Prev Med 14:165–168, 1985
37. Heinonen OP, Albanes D, Virtamo J, Taylor PR, Huttunen JK, Edwards BK, Rautalahti M, Haapakoski J, Pietinen P, Palmgren J, Greenwald P: ATBC lung cancer prevention study. Lung Cancer 7(suppl.):11, 1991

APPENDIX 1

Carotene and Retinol Efficacy Trial Centers and Key Personnel

Baltimore Study Center
Asbestos Cancer Prevention Study
University of Maryland at Baltimore
James Keogh, MD, Principal Investigator
Kate McPhaul, Study Manager

Irvine Study Center
University of California-Irvine
Clinical Cancer Center
Frank Meyskens, Jr., MD, Principal Investigator
James Williams, Jr., MD, Medical Director
Janis DeJohn, Study Manager

New Haven Study Center
CPS/CARET Occupational Health Center
Lawrence & Memorial Hospital
Mark R. Cullen, MD, Principal Investigator
Martin Cherniack, MD, Co-principal Investigator
Brenda Cartmel, PhD, Study Manager

Portland Study Center
Kaiser Permanente Center for Health Research
Barbara Valanis, DrPH, Principal Investigator
Andrew Glass, MD, Co-principal Investigator
Adrianne Feldstein, MD, Co-investigator
Jan Blank, Study Manager

San Francisco Study Center
Asbestos Cancer Prevention Study
San Francisco General Hospital
John Balmes, Principal Investigator
Cynthia Lotane, Study Manager

Seattle Study Center
Fred Hutchinson Cancer Research Center
Gary Goodman, MD, Principal Investigator
Corinne Powell, Study Manager

Coordinating Center
Fred Hutchinson Cancer Research Center
Gilbert S. Omenn, MD, PhD, Principal Investigator
James E. Grizzle, PhD, Director
Bernedine Lund, MS, Study Manager