Practice of Epidemiology

Methods for Pooling Results of Epidemiologic Studies

The Pooling Project of Prospective Studies of Diet and Cancer

Stephanie A. Smith-Warner[1,2], Donna Spiegelman[2,3], John Ritz[2,3], Demetrius Albanes[4], W. Lawrence Beeson[5], Leslie Bernstein[6], Franco Berrino[7], Piet A. van den Brandt[8], Julie E. Buring[2,9], Eunyoung Cho[10], Graham A. Colditz[2,10,11], Aaron R. Folsom[12], Jo L. Freudenheim[13], Edward Giovannucci[1,2,10], R. Alexandra Goldbohm[14], Saxon Graham[13], Lisa Harnack[12], Pamela L. Horn-Ross[15], Vittorio Krogh[7], Michael F. Leitzmann[16], Marjorie L. McCullough[17], Anthony B. Miller[18], Carmen Rodriguez[17], Thomas E. Rohan[19], Arthur Schatzkin[16], Roy Shore[20], Mikko Virtanen[21], Walter C. Willett[1,2,10], Alicja Wolk[22], Anne Zeleniuch-Jacquotte[20], Shumin M. Zhang[2,9], and David J. Hunter[1,2,10,11]

[1] Department of Nutrition, Harvard School of Public Health, Boston, MA.
[2] Department of Epidemiology, Harvard School of Public Health, Boston, MA.
[3] Department of Biostatistics, Harvard School of Public Health, Boston, MA.
[4] Nutritional Epidemiology Branch, National Cancer Institute, Bethesda, MD.
[5] Center for Health Research, School of Medicine, Loma Linda University, Loma Linda, CA.
[6] Department of Preventive Medicine and USC/Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA.
[7] Epidemiology Unit, National Cancer Institute, Milan, Italy.
[8] Department of Epidemiology, Faculty of Health Sciences, Maastricht University, Maastricht, the Netherlands.
[9] Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA.
[10] Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA.
[11] Harvard Center for Cancer Prevention, Boston, MA.
[12] Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN.
[13] Department of Social and Preventive Medicine, University at Buffalo, State University of New York, Buffalo, NY.
[14] Department of Epidemiology, TNO Quality of Life, Zeist, the Netherlands.
[15] Northern California Cancer Center, Fremont, CA.
[16] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.
[17] Epidemiology and Surveillance Research, American Cancer Society, Atlanta, GA.
[18] Department of Public Health Sciences, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada.
[19] Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY.
[20] Department of Environmental Medicine, School of Medicine, New York University, New York, NY.
[21] Department of Epidemiology and Health Promotion, National Public Health Institute, Helsinki, Finland.
[22] Division of Nutritional Epidemiology, National Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden.

With the growing number of epidemiologic publications on the relation between dietary factors and cancer risk, pooled analyses that summarize results from multiple studies are becoming more common. Here, the authors describe the methods being used to summarize data on diet-cancer associations within the ongoing Pooling Project of Prospective Studies of Diet and Cancer, begun in 1991. In the Pooling Project, the primary data from prospective cohort studies meeting prespecified inclusion criteria are analyzed using standardized criteria for modeling of

Correspondence to Dr. Stephanie Smith-Warner, Department of Nutrition, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115 (e-mail: pooling@hsphsun2.harvard.edu).

Am J Epidemiol 2006;163:1053–1064

exposure, confounding, and outcome variables. In addition to evaluating main exposure-disease associations, analyses are also conducted to evaluate whether exposure-disease associations are modified by other dietary and nondietary factors or vary among population subgroups or particular cancer subtypes. Study-specific relative risks are calculated using the Cox proportional hazards model and then pooled using a random- or mixed-effects model. The study-specific estimates are weighted by the inverse of their variances in forming summary estimates. Most of the methods used in the Pooling Project may be adapted for examining associations with dietary and nondietary factors in pooled analyses of case-control studies or case-control and cohort studies combined.

cohort studies; diet; epidemiologic methods; meta-analysis; neoplasms

The growing number of epidemiologic publications on the relation between diet and cancer risk has heightened the need for methods of summarizing results from multiple studies. These methods include qualitative reviews and quantitative summaries such as meta-analyses of the published literature and pooled analyses of the primary data (also called meta-analyses of individual data) (1). A general framework for conducting pooled analyses entails 1) formulating study inclusion criteria; 2) identifying all potential studies meeting these criteria; 3) obtaining each study's primary data; 4) creating a standardized database; 5) estimating study-specific exposure-disease associations; 6) examining whether the study-specific results are heterogeneous; 7) calculating pooled estimates, if applicable; and 8) conducting sensitivity analyses to evaluate whether the estimates are robust (2). There are many advantages to reanalyzing the primary data from multiple studies rather than extracting the study-specific relative risks from published articles (1–5). In a pooled analysis, the modeling of the exposure, confounding, and outcome variables, the choice of which variables to control for, and the type of analysis conducted can be standardized, thereby removing potential sources of heterogeneity across studies. Because of larger sample sizes, pooled analyses also offer investigators the opportunity to examine uncommon exposures, rare diseases, and variation in associations among population subgroups with greater statistical power than is possible in individual studies.

The pooling of data from observational studies has become more common recently (6–13). Summary estimates have been calculated using a weighted average of the study-specific estimates (8, 9, 11) or by combining studies into a single data set for the analysis (6, 7, 10, 12, 13). In this paper, we describe the methods that are being used within the ongoing Pooling Project of Prospective Studies of Diet and Cancer (the Pooling Project), an international consortium of cohort studies with the goal of providing the best available summary of data on associations between diet and cancer (14–30). Most of these methods can also be adapted to examine associations in pooled analyses of case-control studies or both case-control and cohort studies combined.

INCLUSION CRITERIA

To maximize the quality and comparability of the studies in the Pooling Project, we formulated several inclusion criteria a priori. First, we include prospective studies which 1) had at least one publication on the relation between diet and cancer; 2) used a dietary assessment method that was of

sufficient detail to calculate intakes of most nutrients, including energy, and that assessed intake over a period of months or years; and 3) assessed the validity of their dietary assessment method or a closely related instrument. Second, for each cancer site evaluated, we specify a minimum number of cases required for a study to be included in the analysis. Additional inclusion criteria also may be made for each cancer site. Third, for each analysis, we include only those studies that assessed the specified exposure and in which participants consumed the dietary item of interest. For analyses that are going on simultaneously in the Pooling Project and the European Prospective Investigation into Cancer and Nutrition (31), we intend to coordinate analyses so that, to the extent possible, we can use similar analytic approaches and provide comparable results.

COMPONENT STUDIES

Sixteen studies (32–46) are currently included in the Pooling Project (table 1). As we become aware of new studies meeting the inclusion criteria, the investigators from those studies are invited to join the Project. The Canadian National Breast Screening Study and the Netherlands Cohort Study are each analyzed as case-cohort studies (47), because the investigators in these two studies each selected a random sample of the cohort to provide the person-time data for the cohort and have processed questionnaires for only this random sample and the cases. We divide the person-time and numbers of cases compiled during follow-up of the Nurses' Health Study into two segments to take advantage of the expanded food frequency questionnaire administered in 1986 as compared with 1980. In this paper, we refer to the follow-up period from 1980 to 1986 as ''Nurses' Health Study A''; the follow-up period beginning in 1986 is referred to as ''Nurses' Health Study B.'' Following standard survival data analysis theory, blocks of person-time in different time periods are asymptotically uncorrelated, regardless of the extent to which they are derived from the same people (48, 49). Thus, pooling of the estimates from these two time periods produces estimates and standard errors which are as valid as those from a single time period.

Data collection

The investigators in each Pooling Project study send their primary data on select variables to the Harvard School of Public Health (Boston, Massachusetts). There we inspect

the data for completeness and resolve inconsistencies with the investigators of each study.

Each study used a food frequency questionnaire or diet history instrument that was designed and pretested in its specific study population or a similar population (P. L. Horn-Ross, unpublished data; V. Krogh, unpublished data; A. Wolk, unpublished data) (50–59) (table 1). Although the numbers of items included in the food frequency questionnaires varied over fivefold across the studies (table 1), the study-specific correlation coefficients comparing the food frequency questionnaire used in each cohort or a closely related instrument with multiple dietary records or 24-hour recalls generally exceed 0.40 for total fat, dietary fiber, and several micronutrients (P. L. Horn-Ross, unpublished data; V. Krogh, unpublished data; A. Wolk, unpublished data) (50–59) (table 2).

Information on nondietary risk factors was collected at baseline in each study using self-administered questionnaires. For measured covariates, the proportion of missing data for nondietary risk factors is generally low across studies (table 3). The exception is the Swedish Mammography Cohort, in which some covariate information was available for only one of the two counties in the study.

Case ascertainment

Incident cancer diagnoses are identified through follow-up questionnaires, with subsequent medical record review (37, 44, 46), linkage with cancer registries (32, 36, 39–42, 45), or both (33–35, 38, 43). In addition, investigators in some studies ascertain incident and/or fatal outcomes using mortality registries (32, 34, 35, 37–39, 41–46). Case ascertainment has generally been estimated to be greater than 90 percent in each study (table 1).

STATISTICAL APPROACHES AND RATIONALE

For each cohort, after applying the exclusion criteria used in that study, we further exclude participants who reported log_e-transformed energy intakes beyond three standard deviations from the study-specific log_e-transformed mean energy intake of the baseline population (or subcohort, for the case-cohort studies) or who reported a history of cancer (except nonmelanoma skin cancer) at baseline. Additional exclusion criteria may be applied for analyses of specific cancer sites. Because many cancers appear to have hormonal antecedents and because lifestyle factors may differ between women and men, studies including both women and men are split into two studies: a cohort of women and a cohort of men. This conservative approach, in which all estimates are calculated separately for women and men in those studies including both genders, allows for potential effect modification by sex for every determinant of the outcome.

Follow-up time is calculated for each participant from the date on which his/her baseline questionnaire was returned to the date of diagnosis of the specific cancer being examined, the date of death, the date on which the participant moves out of the study area (if applicable), or the end of follow-up, whichever comes first.

In our analyses, we create standardized categories for most confounding variables across studies. We create a missing-data indicator variable for missing responses for each measured confounder in a study, if applicable. As long as 1) the association between the confounding variable and the exposure of interest is weak, or the association between the confounding variable and the outcome is weak, or the confounding variable has little variability in the study and 2) the percentage of missing data within the study is low, the use of the missing-data indicator method is likely to improve efficiency without introducing appreciable bias in comparison with the complete case method (60, 61). As table 3 shows, the proportion of missing data for each covariate across studies is generally low, satisfying one of the conditions for valid use of the missing-data indicator method. In addition, potentially confounding factors generally have had moderate-to-weak associations with the cancer sites we have examined and have had low-to-moderate correlations with the dietary exposures that are of primary interest in the Pooling Project. Information on age, which is typically the strongest measured risk factor for cancer incidence, is never missing in the constituent studies.

Two-stage analysis

Our analytic approach generally is a two-stage process. In the first step, we calculate study-specific relative risks using the Cox proportional hazards model (49), defined through the hazard function h by

$$h_{jks}(t|\boldsymbol{u_{is}}, x_{is}) = h_{0jks}(t)\exp(\boldsymbol{\alpha}_s \boldsymbol{u_{is}} + \beta_s x_{is}) \qquad (1)$$

for $s = 1, \ldots, S$, where s is the study number, t is follow-up time, $\boldsymbol{u_{is}}$ and x_{is} are the study-specific confounding and exposure variables, respectively, for individual i in study s, and $h_{0jks}(t)$ is the baseline incidence rate at age j (in years), in calendar year k, and for time since entry into the study t. The estimated study-specific log relative risks for a one-unit increase in the exposures, x_{is}, are given by the β_s. The study-specific log relative risks for a one-unit increase in the confounding variables, $\boldsymbol{u_{is}}$, are given by the $\boldsymbol{\alpha}_s$. Stratifying jointly by age at baseline (years) and the year in which the baseline questionnaire was returned (indexed by j and k, respectively) and treating follow-up time (in years) as the time metric in the Cox model is equivalent to treating age as the time metric in the Cox model and stratifying jointly on calendar time (in years) and duration of time in the study, with one exception: There is a difference in which two-way interactions are allowed. With our approach, no assumptions are made about the shape of the age or calendar-year incidence curves, each of which is fully adjusted for the other, and arbitrary two-way interactions of the joint dependency of the outcome on age and calendar time are allowed. Each case-cohort study is analyzed using EPICURE software (HiroSoft International Corporation, Seattle, Washington) (47, 62); each remaining study is analyzed using SAS PROC PHREG (SAS Institute, Inc., Cary, North Carolina) (63).

If case-control studies were included in our pooled analyses, the model for these studies would be similar to equation 1, except that we would stratify the participants by

TABLE 1. Characteristics of the studies included in the Pooling Project of Prospective Studies of Diet and Cancer, 1991–2004

Study	Study population	Location	Study dates	Baseline cohort size*		Age (years) at baseline	Food frequency questionnaire/ diet history instrument			Outcome ascertainment	Estimated case ascertainment rate
				Women	Men		No. of items	Time frame	Components measured		
Adventist Health Study (33)	Non-Hispanic White men and women living in Seventh-Day Adventist households	California, United States	1976–1982	18,403	12,896	>24	46	Past year	Frequency	FQs†/MRR†; cancer registry; mortality registry	>99
Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (34)	Male smokers who participated in a randomized double-blind placebo-controlled clinical trial of α-tocopherol and β-carotene supplement use	Southwestern Finland	1985 onward (ongoing)	0	26,987	50–69	276	Past year	Frequency and portion size	FQs/MRR; cancer registry; mortality registry	100
Breast Cancer Detection Demonstration Project Follow-up Cohort (35)	Subset of women participating in a breast cancer screening program in 1973–1980 who had been diagnosed with breast cancer or had undergone or been recommended to receive a breast biopsy, plus a random sample of the remaining women who had been screened	United States	1987 onward (ongoing)	41,987	0	40–93	62	Past year	Frequency and portion size	FQs/MRR; cancer registry; mortality registry	91
California Teachers Study (45)	Active and retired female teachers and administrators participating in the California State Teachers Retirement System	California, United States	1995 onward (ongoing)	100,036	0	21–103	103	Past year	Frequency and portion size	Cancer registry; mortality registry	>97‡
Canadian National Breast Screening Study (36)	Women who participated in a multicenter randomized controlled trial of mammography screening for female breast cancer	Canada	1980 onward (ongoing)	56,837	0	40–59	86	Past month	Frequency and portion size	Cancer registry	100
Cancer Prevention Study II Nutrition Cohort (38)	Subset of men and women participating in Cancer Prevention Study II who completed a diet questionnaire in 1992	United States	1992 onward (ongoing)	74,053	66,090	50–74	68	Past year	Frequency and portion size	FQs/MRR; cancer registry; mortality registry	>90
Health Professionals Follow-up Study (37)	Male dentists, optometrists, osteopathic physicians, podiatrists, pharmacists, and veterinarians	United States	1986 onward (ongoing)	0	47,673	40–75	131	Past year	Frequency of specified portions	FQs/MRR; mortality registry	>94
Iowa Women's Health Study (41)	Postmenopausal women selected randomly from the 1985 Department of Transportation's driver's license list in Iowa	Iowa, United States	1986 onward (ongoing)	34,603	0	55–69	116	Past year	Frequency of specified portions	Cancer registry; mortality registry	98§
Netherlands Cohort Study (40)	Men and women from 204 municipal population registries throughout the Netherlands	The Netherlands	1986 onward (ongoing)	62,573	58,279	55–69	150	Past year	Frequency and portion size	Cancer registry; pathology database	>95

New York State Cohort (42)	Male and female residents who had had the same address and telephone number for the previous 18 years	New York, United States	1980–1987	22,550	30,363	50–93	45	Past year	Frequency	Cancer registry	—¶
New York University Women's Health Study (43)	Women visiting a breast screening clinic who had not used any hormonal medications or been pregnant in the previous 6 months	New York, United States	1985 onward (ongoing)	13,258	0	34–65	71	Past year	Frequency and portion size	FQs/MRR; cancer registry; mortality registry	95
Nurses' Health Study A (37)	Female registered nurses	United States	1980–1986	88,651	0	34–59	61	Past year	Frequency of specified portions	FQs/MRR; mortality registry	>94
Nurses' Health Study B (37)	Female registered nurses	United States	1986 onward (ongoing)	68,540	0	40–65	131	Past year	Frequency of specified portions	FQs/MRR; mortality registry	>94
Nurses' Health Study II (46)	Female registered nurses	United States	1991 onward (ongoing)	93,894	0	26–46	133	Past year	Frequency of specified portions	FQs/MRR; mortality registry	>90
Prospective Study on Hormones, Diet and Breast Cancer (39)	Female volunteers recruited from the general population using mass media advertising and from breast cancer prevention units	Varese Province, Italy	1987 onward (ongoing)	9,027	0	35–69	177	Past year	Frequency and portion size	Cancer registry; mortality registry; admissions and discharge reports; pathology database	>97
Swedish Mammography Cohort (32)	Women who participated in a population-based mammography screening program	Västmanland and Uppsala counties, Sweden	1987 onward (ongoing)	61,463	0	40–74	67	Past 6 months	Frequency	Cancer registry	98
Women's Health Study (44)	Female health professionals who participated in a randomized, double-blind, placebo-controlled trial of low-dose aspirin, β-carotene, and vitamin E use	United States	1993 onward (ongoing)	38,384	0	≥45	131	Past year	Frequency of specified portions	FQs/MRR	96

* The baseline cohort size corresponds to the number of participants in the Pooling Project database for the renal cell cancer analyses in the California Teachers Study (45) and for the colorectal cancer analyses in the remaining studies.

† FQs, follow-up questionnaires; MRR, medical record review.

‡ For California residents only.

§ For Iowa residents only.

¶ Cancer outcomes in the New York State Cohort (42) were identified through linkage with a cancer registry; thus, it is difficult to determine the follow-up rate in the cohort. When a subset of the cohort was followed intensively, loss to follow-up was not related to exposure.

TABLE 2. Correlation coefficients (CCs) for nutrient intakes estimated using a food frequency questionnaire versus a comparison method for studies in the Pooling Project of Prospective Studies of Diet and Cancer, 1991–2004*

Study	Sex	No. of participants	Comparison method	Type of CC	Total fat	Saturated fat	Mono-unsaturated fat	Poly-unsaturated fat	Dietary fiber	Alcohol	Vitamin A†	Vitamin C†	Vitamin E†	Folate†	Calcium†
Adventist Health Study (50)	Women	103	Five 24-hour recalls over 6 months	Spearman CCs‡	0.40	0.45	0.41	0.26	0.47§						
	Men	44	Five 24-hour recalls over 6 months	Spearman CCs‡	0.38	0.57	0.29	0.15	0.50§						
Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (51)	Men	178	Twelve 2-day diet records over 6 months	Energy-adjusted, deattenuated Pearson CCs	0.75	0.79	0.68	0.85	0.82	0.85	0.68	0.71	0.82		0.74
California Teachers Study (unpublished)	Women	185	Four 24-hour recalls over 10 months	Energy-adjusted, deattenuated Pearson CCs	0.64	0.82	0.41	0.23	0.77	0.82	0.35¶	0.62¶	0.82¶	0.73¶	0.30¶
Canadian National Breast Screening Study (52)	Women	108	7-day diet records	Energy-adjusted Pearson CCs	0.44	0.61	0.43#	0.40‡‡	0.60		0.59	0.60	0.53		0.67
Cancer Prevention Study II Nutrition Cohort (53)	Women	188‡	Four 24-hour recalls over 1 year	Energy-adjusted, deattenuated Pearson CCs	0.66	0.66	0.58#	0.42‡‡	0.61	0.77		0.65	0.27	0.43	0.66
	Men	229‡	Four 24-hour recalls over 1 year	Energy-adjusted, deattenuated Pearson CCs	0.58	0.64	0.61#	0.48‡‡	0.64	0.82		0.65	0.23	0.51	0.57
Health Professionals Follow-up Study (54)	Men	127	Two 7-day diet records over 6 months	Energy-adjusted, deattenuated Pearson CCs	0.67	0.75	0.68	0.37	0.68	0.86‡,**	0.61	0.77	0.42	0.70	0.60
Iowa Women's Health Study (55)	Women	44	Five 24-hour recalls over 2 months	Energy-adjusted Pearson CCs	0.62	0.59	0.62	0.43	0.24§	0.32	0.14	0.53	0.79	0.26	0.49
Netherlands Cohort Study (56)	Women and men	109	Three 3-day diet records over 1 year	Energy- and sex-adjusted deattenuated Pearson CCs	0.53	0.58††		0.80	0.79	0.86	0.76	0.58			0.66
New York State Cohort	Women (unpublished)	190	Simulated study	Energy-adjusted, deattenuated Pearson CCs‡	0.21	0.18			0.41		0.25	0.53	0.16		
	Men (57)	127	Simulated study	Energy-adjusted, deattenuated Pearson CCs	0.57	0.60	0.61	0.22	0.65		0.39	0.76		0.46	0.60
Nurses' Health Study A (58)	Women	173	Four 7-day diet records over 1 year	Energy-adjusted Pearson CCs	0.53	0.59		0.48	0.58§	0.90‡,**	0.36	0.66			
Nurses' Health Study B (59)	Women	191	Two 7-day diet records over 1 year	Energy-adjusted, deattenuated Pearson CCs	0.57	0.68	0.58	0.48				0.79	0.76		0.75

Prospective Study on Hormones, Diet and Breast Cancer (unpublished)	Women	104	Spearman CCs	0.39	0.49	0.40	0.42	0.53	0.88	0.19	0.37	0.37	0.32	0.48
Swedish Mammography Cohort (unpublished)	Women	129	Energy-adjusted, deattenuated Pearson CCs	0.49	0.42	0.51	0.36	0.54	0.85	0.49	0.33	0.29	0.48	0.48
Median				0.55	0.60	0.55	0.40	0.64	0.84	0.45	0.65	0.37	0.46	0.60

* A blank cell means that the investigators did not evaluate this nutrient in their validation study. The studies not mentioned in this table used questionnaires that were very similar to those that had been validated previously by other investigators. The Breast Cancer Detection Demonstration Project Follow-up Cohort (35) and the New York University Women's Health Study (43) both used food frequency questionnaires similar to the food frequency questionnaire used in the Cancer Prevention Study II Nutrition Cohort (53). Nurses' Health Study II (46) and the Women's Health Study (44) both used food frequency questionnaires similar to the food frequency questionnaire used in Nurses' Health Study B (59).

† Intake from foods only; supplemental intake was not included.

‡ The data presented were calculated using the validation study data that were sent to the Pooling Project.

§ Crude fiber.

¶ Correlations among nonusers of supplements only (n = 44).

Oleic acid.

** Spearman CC.

†† Energy- and sex-adjusted Pearson CC.

‡‡ Linoleic acid.

study-specific matching factors. Since the Cox model and the conditional logistic regression model produce algebraically identical log-(partial)-likelihood functions, SAS PROC PHREG could also be used for case-control studies to estimate study-specific odd ratios and their standard errors.

The second step consists of pooling the study-specific relative risks using a random-effects model (64–66) given by

$$\hat{\beta}_s = \beta + b_s + e_s, \qquad (2)$$

where the $\hat{\beta}_s$ are the estimated study-specific exposure-disease effects, β is the underlying common exposure-disease association, b_s are the random between-studies effects, and e_s are the within-study errors. Both b_s and e_s are assumed to be independent and asymptotically normally distributed with means of zero and variances of σ_B^2 and σ_s^2, respectively, and $\sigma_s^2 = \mathrm{Var}(\hat{\beta}_s)$. The study-specific exposure-disease effects are weighted by the inverse of their variances using

$$\hat{\beta} = \sum_{s=1}^{S} w_s \hat{\beta}_s,$$

where

$$w_s = (\hat{\sigma}_B^2 + \hat{\sigma}_s^2)^{-1} \Big/ \sum_{t=1}^{S} (\hat{\sigma}_B^2 + \hat{\sigma}_t^2)^{-1}.$$

When the exposure variable is categorized into different levels, we calculate a pooled relative risk for each category separately.

We test for the statistical significance of between-studies heterogeneity among the study-specific exposure-disease estimates using the Q test statistic given by

$$Q = \sum_{s=1}^{S} w_s^* (\hat{\beta}_s - \hat{\beta})^2, \qquad (3)$$

where $w_s^* = \widehat{\mathrm{Var}}(\hat{\beta}_s)^{-1}$. The Q test statistic follows an approximate χ_{s-1}^2 distribution (66, 67).

For the exposures of interest, we generally categorize participants into study-specific quantiles. Because the quantile approach does not take into account true differences in the distribution of population intakes across studies, we also create categories defined by identical absolute intake cutpoints across studies. Misclassification can also occur in the analyses based on identical absolute intake cutpoints, because reported intakes may differ across studies based on differences in the dietary assessment methods used. However, when possible we adjust our results for measurement error in the individual studies.

Aggregated analysis

We can also conduct analyses in which the data from all studies are combined into one data set (referred to as an aggregated analysis). A single exposure-disease effect is then calculated using the Cox proportional hazards model, including stratification by study, age at baseline, and the year in which the baseline questionnaire was returned.

TABLE 3. Prevalences of missing data for select nondietary factors across studies in the Pooling Project of Prospective Studies of Diet and Cancer, 1991–2004

	No. of studies in which the factor was measured*	% of missing data across studies (range)	Studies with <5% missing data	
			No.	%
Age	17	0	17	100
Education	17†	0–23	14	82
Body mass index	17	0–8	15	88
Smoking status	15	0–5	15	100
Physical activity	14	0–12	11	79
Multivitamin use	14	0–8	10	71
Age at menarche‡,§	14	0–3	14	100
Parity‡	15	0–10	13	87
Menopausal status‡,§	14	0–18	8	57
Oral contraceptive use‡,§	13	0–21	11	85
Postmenopausal hormone use‡,§,¶	13	0–16	12	92

* For this table, Nurses' Health Study A (1980–1986) and Nurses' Health Study B (1986–present) were counted as two separate studies (see Materials and Methods).

† All participants in the California Teachers Study, the Health Professionals Follow-up Study, the Nurses' Health Study, and Nurses' Health Study II were assumed to have received additional education after graduating from high school, because these populations were selected on the basis of their employment in occupations requiring a post-high-school education.

‡ Only cohort studies including women are included here. The prevalence of missing data was calculated only among the female participants.

§ For the Swedish Mammography Cohort, only the percentage of missing data for women living in Uppsala County is included, since these data were not collected for women living in Vastmanland County.

¶ Among postmenopausal women only.

Although combining the data from all studies is one way to take advantage of differences in the distributions of the exposure variable across studies, it assumes that the exposure was measured in comparable ways across studies. Because the distributions of dietary variables may differ across studies due to true differences in actual intake and due to differences in the dietary assessment methods used (and other study-specific sources of error), this assumption may not be reasonable, except for nutrients that come from a small number of food sources (e.g., alcohol). In addition, combining the studies into one data set assumes that there is no between-studies heterogeneity in the associations of the outcome with the exposure or any of the covariates. In the few instances where we have conducted both pooled and aggregated analyses, the results have been essentially identical (16, 25, 30). Nevertheless, because it is difficult to test the underlying assumptions, we have opted to use two-stage analyses as our primary analytic strategy.

Trend analysis

To test the significance of trends in disease risk over exposure categories, we conduct separate analyses in which participants are assigned the study-specific median value of their respective category (given by med_{js} for $j = 1, \ldots, J$, where J is the number of levels in which the exposure variable is categorized). For each study, we fit a Cox proportional

hazards model with regression terms $\beta_s z_{is}$ for $s = 1, \ldots, S$, where s is the study number and z_{is} takes on the values med_{js} corresponding to the category in which the individual's exposure value falls. We then compute the pooled estimate for the regression coefficient for trend using a random-effects model (64–66). The pooled test for trend is a Wald test of the hypothesis H_0: $\beta = 0$. We test for the statistical significance of between-studies heterogeneity among the study-specific regression coefficients using the Q test statistic (66, 67).

We also evaluate whether associations between dietary factors and cancer risk are linear by comparing nonparametric regression curves using restricted cubic splines with the linear model using the likelihood ratio test, and by visual inspection of the restricted cubic spline graphs (68, 69). For these analyses, the studies are combined into a single data set stratified by study.

Evaluation of heterogeneity of effects

An advantage of a pooled analysis is the ability to evaluate whether the exposure-disease association is modified by other risk factors. In these analyses, if the exposure-disease association is log-linear and the potential effect modifier is an ordinal or binary variable, we first compute estimates of the exposure-disease association and their standard errors for each study within each category of the

potential effect modifier. The model uses the same format as equation 1, but

$$h_{jksl}(t|\boldsymbol{u_{is}},x_{is}) = h_{0jksl}(t)\exp(\boldsymbol{\alpha}_{sl}\boldsymbol{u_{is}} + \beta_{sl}x_{is}), \qquad (4)$$

where $l = 1, \ldots, L$ levels of the effect modifier, x_{is} is the study-specific exposure variable, and $\boldsymbol{\alpha}_{sl}$ are the estimated study-specific log relative risks for a one-unit increase in the confounding variables, $\boldsymbol{u_{is}}$. The study-specific estimates $\hat{\beta}_{sl}$ for each stratum are then pooled across studies and exponentiated to obtain the relative risk for each level of the potential effect modifier. For assessment of the statistical significance of the interaction, the Cox proportional hazards model is

$$\begin{aligned} h_{jks}(t|\boldsymbol{u_{is}},x_{is},m_{is}) = h_{0jks}(t)\exp(\boldsymbol{\alpha}_s\boldsymbol{u_{is}} + \xi_s m_{is} \\ + \beta_s x_{is} + \gamma_s x_{is}m_{is}), \end{aligned} \qquad (5)$$

where γ_s is the study-specific estimate for the cross-product term of the potential effect modifier variable (m_{is}) times the exposure variable (x_{is}) and ξ_s is the study-specific main effect of the effect modifier. The study-specific estimates $\hat{\gamma}_s$ are then pooled across studies, and the p value corresponding to the test for interaction ($H_0: \gamma = 0$) is obtained from a Wald test based upon the pooled $\hat{\gamma}$.

We use a mixed-effects meta-regression model (70) to test for effect modification when the exposure-disease association is nonlinear, when the potential effect modifier is a polytomous nominal variable, or when effect modification can be assessed only between studies. As an example, consider the test for effect modification by gender. The model here is a slightly modified version of equation 2:

$$\hat{\beta}_s = \beta_0 + \beta_1 z_s + b_s + e_s \qquad (6)$$

for $s = 1, \ldots, S$, where s is the study number, the $\hat{\beta}_s$ are the estimated study-specific exposure-disease effects, β_0 is the log relative risk for the exposure in the reference level of the modifier (here, men), β_1 is the difference in the log relative risks between the reference level and each of the other levels (here, between genders), $z_s = 1$ if study s is carried out among women and $z_s = 0$ if it is carried out among men, b_s are the study-specific random effects, and e_s are the within-study sampling errors. The Wald test statistic based on the estimate $\hat{\beta}_1$ and its standard error is used to test the null hypothesis ($H_0: \beta_1 = 0$) that there is no modification of the effect of exposure on the outcome by levels of the potential effect modifier (here, between genders).

Assessment of heterogeneity by outcome subtype

We can also evaluate whether associations differ by cancer subtype. For these analyses, we fit separate Cox proportional hazards models (equation 1) for each subtype. Occurrences of the cancer under study that are of a different subtype are censored at their date of diagnosis. The relative risks obtained for each subtype that are estimated in this way are asymptotically uncorrelated (71–73). In addition, because these estimates are asymptotically normally distributed with variances given by the square of their respective estimated standard errors, any linear combination of the

different estimates is normally distributed, and it follows from the Cramer-Wald device (74) that the multivariate vector obtained by combining all of the competing risk estimates is multivariate normal. The corresponding variances are in the diagonal of the covariance matrix, and zeroes are in the off-diagonal. To test the null hypothesis that there is no difference in the pooled exposure-disease parameters among the subtypes, we use a contrast test (75). For example, to test whether the pooled exposure-disease parameters differed among three subtypes, we would use the test statistic Z^2 given by

$$Z^2 = (\boldsymbol{C}\hat{\boldsymbol{\beta}})^T(\boldsymbol{C}\hat{\Sigma}\boldsymbol{C}^T)^{-1}(\boldsymbol{C}\hat{\boldsymbol{\beta}}), \qquad (7)$$

where C is a contrast matrix whose first and second rows are $(1, -1, 0)$ and $(1, 0, -1)$, $\hat{\beta}$ is the vector of the pooled estimates of the exposure-disease association for the different subtypes, and $\hat{\Sigma}$ is its estimated covariance matrix. The Z^2 statistic in this example has an approximate χ^2 distribution with 2 df (defined by the number of different subtypes minus 1) (75). These methods can also be used to construct tests for heterogeneity of effects between any set of cancers or other outcomes.

Measurement error correction

As with most exposures, measurement of dietary variables is not free from error. Measurement error in dietary data derives from normal within-person variation in intakes over time (76) and from errors associated with self-reports (77). Therefore, the relative risks will be biased, usually towards the null, but can be biased in either direction when there is also error in measuring confounding variables (78). One can use the validation data from each study to regress the "gold standard" (or an unbiased estimate of the gold standard, an "alloyed" gold standard (79)) on the error-prone measurement and confounding variables to obtain a correction factor. This correction factor can then be used to calibrate the uncorrected estimates of the exposure effect of interest obtained from logistic and Cox regression models (77, 79, 80). If the errors in the alloyed gold standard are correlated with the errors in the usual measure of dietary intake, the regression calibration method for measurement error correction will remove some, but not all, of the bias in the effect estimate (81). However, it appears that energy adjustment removes much of the bias in this method due to correlated errors for at least some dietary variables (e.g., protein) (82, 83). To remove the remaining bias, an additional method of assessment of intake is needed, such as a biomarker (81).

In the measurement error correction analyses, for each study, the true intake of the particular nutrient being evaluated or an unbiased estimate of the true intake (e.g., intakes calculated from several dietary records or 24-hour recalls) is regressed on the surrogate measurement of that nutrient (calculated from the food frequency questionnaire) to obtain the coefficient $\hat{\lambda}_s$ and its estimated standard error. We then derive the corrected estimate of the log relative risk as $\hat{\beta}_s/\hat{\lambda}_s$, where $\hat{\beta}_s$ is the uncorrected estimated effect in each study from a logistic regression or Cox proportional hazards

regression analysis. The standard error of $\hat{\beta}_s/\hat{\lambda}_s$ is derived using the delta method (84). One can simultaneously correct for the error in several covariates in all point estimates and their standard errors using a multivariate extension of measurement error correction (79, 85). The corrected coefficient estimates are then pooled into a summary estimate. If a study has poor validity of nutrient measurements, its variance will be large, and the study will thus have little weight when the study-specific results are pooled. In addition, under the required assumption that the dietary records and 24-hour recalls provide an unbiased estimate of nutrient intake (even if subject to random error), this approach calibrates the estimated relative risks to a common unit of measurement across studies, thereby adjusting for systematic errors due to differences in the food frequency questionnaires used in the various studies.

STRENGTHS AND LIMITATIONS

The Pooling Project of Prospective Studies of Diet and Cancer provides a large collection of data in which multiple diet-and-cancer hypotheses can be examined with greater statistical power than is available in any one study. Each study included in the Pooling Project is a prospective cohort study in which diet was assessed prior to development of disease, thereby limiting recall and selection biases. In the Pooling Project, we standardize the modeling of the exposure and confounding variables to remove potential sources of noncomparability and heterogeneity that occur in the published literature. We are able to examine associations over a wide range of intakes with greater precision than in the individual studies, because of the larger sample size and the different diets consumed across the populations. In addition, we can evaluate whether associations are modified by other factors and whether associations differ among cancer subtypes. Because inclusion of an individual study in a particular analysis is not dependent on whether those investigators have published findings on that association, publication bias does not affect our pooled analyses—as opposed to meta-analyses of the published literature, for which approximately half of the results may have some indication of publication bias (86). Finally, results from these pooled analyses may assist epidemiologists and other health professionals in synthesizing the vast amount of published data on specific diet-cancer associations.

A limitation of the Pooling Project is that it was planned retrospectively. Thus, there are differences in how the included studies were designed and implemented. First, the studies comprise populations from different geographic regions with different age ranges and education levels. However, these differences in study population characteristics may be considered a strength, particularly if the results are consistent across studies. Second, the dietary assessment methods used vary across studies, which may lead to artifactual differences in estimated intakes across studies, in addition to any true between-population differences in intakes. However, it is also possible that validity is enhanced by the use of study-specific questionnaires, since they may be tailored for use in each component study. Some heterogeneity of assessment instruments cannot be avoided, even in prospectively planned pooled studies—if, for instance, the language spoken and the foods consumed differ between populations. Another limitation of the Pooling Project is that only current diet at baseline was measured in most of the studies; thus, we cannot examine the effects of dietary changes occurring during follow-up or assess associations with diet at younger ages. There may be differential control for confounding across studies because the nondietary variables that were measured varied across studies, although many important potential confounders were measured in most studies. In addition, by standardizing which confounding variables are included in the multivariate models and their categorization, we have minimized between-studies heterogeneity resulting from how potentially confounding variables were modeled. A final restriction is our inability to examine effect modification by race and ethnicity, because the Pooling Project currently includes studies from only North America and Europe and a predominantly Caucasian population; however, as studies from other regions and with persons of different ethnicities become eligible to join the Pooling Project, the ethnic composition of the Pooling Project will expand.

Despite these limitations and restrictions, the data compiled in the Pooling Project are a valuable resource for prospectively investigating associations between diet and cancer, particularly for population subgroups, less common cancers, and specific cancer subtypes. In our analyses, we use standardized criteria to define each variable in order to reduce potential sources of between-studies heterogeneity. We then evaluate whether associations are consistent across different study populations. Finally, the methods that we use in the Pooling Project may be modified to pool data from both case-control and cohort studies to examine associations between dietary and nondietary risk factors and disease.

REFERENCES

1. Blettner M, Sauerbrei W, Schlehofer B, et al. Traditional reviews, meta-analyses and pooled analyses in epidemiology. Int J Epidemiol 1999;28:1–9.
2. Friedenreich CM. Methods for pooled analyses of epidemiologic studies. Epidemiology 1993;4:295–302.
3. Steinberg KK, Smith SJ, Stroup DF, et al. Comparison of effect estimates from a meta-analysis of summary data from published studies and from a meta-analysis using individual patient data for ovarian cancer studies. Am J Epidemiol 1997; 145:917–25.
4. Lyman GH, Kuderer NM. The strengths and limitations of meta-analyses based on aggregate data. BMC Med Res Methodol 2005;5:14.
5. Ioannidis JP, Rosenberg PS, Goedert JJ, et al. Commentary: meta-analysis of individual participants' data in genetic epidemiology. Am J Epidemiol 2002;156:204–10.

6. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormonal contraceptives: collaborative re-analysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. Lancet 1996;347:1713–27.

7. Whittemore AS, Harris R, Itnyre J, et al. Characteristics relating to ovarian cancer risk: collaborative analysis of 12 US case-control studies. I. Methods. Am J Epidemiol 1992; 136:1175–83.

8. Plummer M, Herrero R, Franceschi S, et al. Smoking and cervical cancer: pooled analysis of the IARC multi-centric case-control study. Cancer Causes Control 2003;14:805–14.

9. Bosetti C, Kolonel L, Negri E, et al. A pooled analysis of case-control studies of thyroid cancer. VI. Fish and shellfish consumption. Cancer Causes Control 2001;12:375–82.

10. Arslan AA, Zeleniuch-Jacquotte A, Lundin E, et al. Serum follicle-stimulating hormone and risk of epithelial ovarian cancer in postmenopausal women. Cancer Epidemiol Biomarkers Prev 2003;12:1531–5.

11. Pereira MA, O'Reilly E, Augustsson K, et al. Dietary fiber and risk of coronary heart disease: a pooled analysis of cohort studies. Arch Intern Med 2004;164:370–6.

12. Morton LM, Hartge P, Holford TR, et al. Cigarette smoking and risk of non-Hodgkin lymphoma: a pooled analysis from the International Lymphoma Epidemiology Consortium (InterLymph). Cancer Epidemiol Biomarkers Prev 2005;14: 925–33.

13. Smith JS, Herrero R, Bosetti C, et al. Herpes simplex virus-2 as a human papillomavirus cofactor in the etiology of invasive cervical cancer. J Natl Cancer Inst 2002;94:1604–13.

14. Hunter DJ, Spiegelman D, Adami H-O, et al. Cohort studies of fat intake and the risk of breast cancer—a pooled analysis. N Engl J Med 1996;334:356–61.

15. Hunter DJ, Spiegelman D, Adami H-O, et al. Non-dietary factors as risk factors for breast cancer, and as effect modifiers of the association of fat intake and risk of breast cancer. Cancer Causes Control 1997;8:49–56.

16. Smith-Warner SA, Spiegelman D, Yaun S-S, et al. Alcohol and breast cancer in women: a pooled analysis of cohort studies. JAMA 1998;279:535–40.

17. Cho E, Smith-Warner SA, Spiegelman D, et al. Dairy foods, calcium, and colorectal cancer: a pooled analysis of 10 cohort studies. J Natl Cancer Inst 2004;96:1015–22.

18. van den Brandt PA, Spiegelman D, Yaun SS, et al. Pooled analysis of prospective cohort studies on height, weight and breast cancer risk. Am J Epidemiol 2000;152:514–27.

19. Smith-Warner SA, Spiegelman D, Yaun S-S, et al. Intake of fruits and vegetables and risk of breast cancer: a pooled analysis of cohort studies. JAMA 2001;285:769–76.

20. Smith-Warner SA, Spiegelman D, Adami HO, et al. Types of dietary fat and breast cancer: a pooled analysis of cohort studies. Int J Cancer 2001;92:767–74.

21. Missmer SA, Smith-Warner SA, Spiegelman D, et al. Meat and dairy food consumption and breast cancer: a pooled analysis of cohort studies. Int J Epidemiol 2002;31:78–85.

22. Smith-Warner SA, Ritz J, Hunter DJ, et al. Dietary fat and risk of lung cancer in a pooled analysis of prospective studies. Cancer Epidemiol Biomarkers Prev 2002;11:987–92.

23. Smith-Warner SA, Spiegelman D, Yaun SS, et al. Fruits, vegetables and lung cancer: a pooled analysis of cohort studies. Int J Cancer 2003;107:1001–11.

24. Mannisto S, Smith-Warner SA, Spiegelman D, et al. Dietary carotenoids and risk of lung cancer in a pooled analysis of seven cohort studies. Cancer Epidemiol Biomarkers Prev 2004;13:40–8.

25. Cho E, Smith-Warner SA, Ritz J, et al. Alcohol intake and colorectal cancer: a pooled analysis of 8 cohort studies. Ann Intern Med 2004;140:603–13.

26. Koushik A, Hunter DJ, Spiegelman D, et al. Fruits and vegetables and ovarian cancer risk in a pooled analysis of 12 cohort studies. Cancer Epidemiol Biomarkers Prev 2005;14: 2160–7.

27. Freudenheim JL, Ritz J, Smith-Warner SA, et al. Alcohol consumption and risk of lung cancer: a pooled analysis of cohort studies. Am J Clin Nutr 2005;82:657–67.

28. Cho E, Hunter DJ, Spiegelman D, et al. Intakes of vitamins A, C and E and folate and multivitamins and lung cancer: a pooled analysis of 8 prospective studies. Int J Cancer 2006; 118:970–8.

29. Genkinger JM, Hunter DJ, Spiegelman D, et al. A pooled analysis of 12 cohort studies of dietary fat, cholesterol and egg intake and ovarian cancer. Cancer Causes Control 2006;17: 273–85.

30. Park Y, Hunter DJ, Spiegelman D, et al. Dietary fiber intake and risk of colorectal cancer: a pooled analysis of prospective cohort studies. JAMA 2005;294:2849–57.

31. Riboli E, Kaaks R. The EPIC Project: rationale and study design. Int J Epidemiol 1997;26(suppl 1):S6–14.

32. Wolk A, Bergström R, Hunter D, et al. A prospective study of association of monounsaturated fat and other types of fat with risk of breast cancer. Arch Intern Med 1998;158: 41–5.

33. Singh PN, Fraser GE. Dietary risk factors for colon cancer in a low-risk population. Am J Epidemiol 1998;148:761–74.

34. 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.

35. Flood A, Velie EM, Chaterjee N, et al. Fruit and vegetable intakes and the risk of colorectal cancer in the Breast Cancer Detection Demonstration Project Follow-up Cohort. Am J Clin Nutr 2002;75:936–43.

36. Terry P, Jain M, Miller AB, et al. Dietary intake of folic acid and colorectal cancer risk in a cohort of women. Int J Cancer 2002;97:864–7.

37. Michels KB, Giovannucci E, Joshipura KJ, et al. Prospective study of fruit and vegetable consumption and incidence of colon and rectal cancers. J Natl Cancer Inst 2000;92: 1740–52.

38. Calle EE, Rodriguez C, Jacobs EJ, et al. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. Cancer 2002; 94:500–11.

39. Sieri S, Krogh V, Muti P, et al. Fat and protein intake and subsequent breast cancer risk in postmenopausal women. Nutr Cancer 2002;42:10–17.

40. Voorrips LE, Goldbohm RA, van Poppel G, et al. Vegetable and fruit consumption and risks of colon and rectal cancer in a prospective cohort study: The Netherlands Cohort Study on Diet and Cancer. Am J Epidemiol 2000;152:1081–92.

41. Steinmetz KA, Kushi LH, Bostick RM, et al. Vegetables, fruit, and colon cancer in the Iowa Women's Health Study. Am J Epidemiol 1994;139:1–15.

42. Bandera EV, Freudenheim JL, Marshall JR, et al. Diet and alcohol consumption and lung cancer risk in the New York State Cohort (United States). Cancer Causes Control 1997; 8:828–40.

43. Kato I, Akhmedkhanov A, Koenig K, et al. Prospective study of diet and female colorectal cancer: The New York University Women's Health Study. Nutr Cancer 1997;28:276–81.

44. Higginbotham S, Zhang Z-F, Lee I-M, et al. Dietary glycemic load and risk of colorectal cancer in the Women's Health Study. J Natl Cancer Inst 2004;96:229–33.

45. Horn-Ross PL, Hoggatt KJ, West DW, et al. Recent diet and breast cancer risk: The California Teachers Study (USA). Cancer Causes Control 2002;13:407–15.

46. Cho E, Spiegelman D, Hunter DJ, et al. Premenopausal fat intake and risk of breast cancer. J Natl Cancer Inst 2003;95: 1079–85.

47. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika 1986;73: 1–11.

48. Rothman KJ. Modern epidemiology. Boston, MA: Little, Brown and Company, 1986.

49. Cox DR. Regression models and life tables (with discussion). J R Stat Soc B 1972;34:187–220.

50. Abbey DE, Andress M, Fraser G, et al. Validity and reliability of alternative nutrient indices based on a food frequency questionnaire. (Abstract). Am J Epidemiol 1988;128(suppl): 934.

51. Pietinen P, Hartman AM, Haapa E, et al. Reproducibility and validity of dietary assessment instruments. I. A self-administered food use questionnaire with a portion size picture booklet. Am J Epidemiol 1988;128:655–66.

52. Jain M, Howe GR, Rohan T. Dietary assessment in epidemiology: comparison of a food frequency and a diet history questionnaire with a 7-day food record. Am J Epidemiol 1996;143:953–60.

53. Flagg EW, Coates RJ, Calle EE, et al. Validation of the American Cancer Society Cancer Prevention Study II Nutrition Survey Cohort food frequency questionnaire. Epidemiology 2000;11:462–8.

54. Rimm EB, Giovannucci EL, Stampfer MJ, et al. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am J Epidemiol 1992;135:1114–26.

55. Munger RG, Folsom AR, Kushi LH, et al. Dietary assessment of older Iowa women with a food frequency questionnaire: nutrient intake, reproducibility, and comparison with 24-hour dietary recall interviews. Am J Epidemiol 1992;136:192–200.

56. Goldbohm RA, van den Brandt PA, Brants HA, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr 1994; 48:253–65.

57. Feskanich D, Marshall J, Rimm EB, et al. Simulated validation of a brief food frequency questionnaire. Ann Epidemiol 1994; 4:181–7.

58. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122:51–65.

59. Willett W. Nutritional epidemiology. New York, NY: Oxford University Press, 1998.

60. Miettinen OS. Theoretical epidemiology. New York, NY: John Wiley and Sons, Inc, 1985.

61. Huberman M, Langholz B. Application of the missing-indicator method in matched case-control studies with incomplete data. Am J Epidemiol 1999;150:1340–5.

62. HiroSoft International Corporation. EPICURE user's guide: the PEANUTS program. Seattle, WA: HiroSoft International Corporation, 1993.

63. Allison PD. Survival analysis using the SAS system: a practical guide. Cary, NC: SAS Publishing, 1995.

64. Harville DA. Maximum likelihood approaches to variance component estimation and to related problems. J Am Stat Assoc 1977;72:320–38.

65. Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982;38:963–74.

66. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.

67. Cochran WG. The combination of estimates from different experiments. Biometrics 1954;10:101–29.

68. Durrleman S, Simon R. Flexible regression models with cubic splines. Stat Med 1989;8:551–61.

69. Smith PL. Splines as a useful and convenient statistical tool. Am Stat 1979;33:57–62.

70. Stram DO. Meta-analysis of published data using a linear mixed-effects model. Biometrics 1996;52:536–44.

71. Prentice RL, Kalbfleisch JD, Peterson AV, et al. The analysis of failure times in the presence of competing risks. Biometrics 1978;34:541–54.

72. Tsiatis AA. Competing risks. In: Armitage P, Colton T, eds. Encyclopedia of biostatistics. 1st ed. Vol 1. New York, NY: John Wiley and Sons, Inc, 1988:824–34.

73. Cox DR, Oakes D. Analysis of survival data. New York, NY: Chapman and Hall, Inc, 1993.

74. Billingsley P. Probability and measure. New York, NY: John Wiley and Sons, Inc, 1995.

75. Anderson TW. Introduction to multivariate statistics. New York, NY: John Wiley and Sons, Inc, 1984.

76. Beaton GH, Milner J, McGuire V, et al. Source of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr 1983;37:986–95.

77. Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. Stat Med 1989; 8:1051–69.

78. Kupper LL. Effects of the use of unreliable surrogate variables on the validity of epidemiologic research studies. Am J Epidemiol 1984;120:643–8.

79. Spiegelman D, Schneeweiss S, McDermott A. Measurement error correction for logistic regression models with an "alloyed gold standard." Am J Epidemiol 1997;145:184–96.

80. Wang CY, Xie, Prentice AM, et al. Recalibration based on an approximate relative risk estimator in Cox regression with missing covariates. Stat Sinica 2001;11:1081–104.

81. Spiegelman D, Zhao B, Kim J. Correlated errors in biased surrogates: study designs and methods for measurement error correction. Stat Med 2005;24:1657–82.

82. Kipnis V, Subar AF, Midthune D, et al. Structure of dietary measurement error: results of the OPEN biomarker study. Am J Epidemiol 2003;158:14–21.

83. Michels KB, Bingham SA, Luben R, et al. The effect of correlated measurement error in multivariate models of diet. Am J Epidemiol 2004;160:59–67.

84. Bishop Y, Fienberg S, Holland P. Discrete multivariate analysis. Cambridge, MA: MIT Press, 1975.

85. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. Am J Epidemiol 1990;132:734–45.

86. Sutton AJ, Duval SJ, Tweedie RL, et al. Empirical assessment of effect of publication bias on meta-analyses. BMJ 2000; 320:1574–7.