SUPPLEMENTAL METHODS

*Colon Cancer Family Registry (CCFR)*

The CCFR is a National Cancer Institute–supported consortium consisting of 6 centers dedicated to the establishment of a comprehensive collaborative infrastructure for interdisciplinary studies in the genetic epidemiology of colorectal cancer [1]. The CCFR includes data from approximately 30,500 total subjects (10,500 probands and 20,000 unaffected and affected relatives and unrelated controls). Cases and controls, age 20–74 years, were recruited from the 6 participating centers beginning in 1998. CCFR implemented a standardized questionnaire that was administered to all participants; this questionnaire included established and suspected risk factors for colorectal cancer, including questions on medical history and medication use, reproductive history (for female participants), family history, physical activity, demographics, alcohol and tobacco use, and dietary factors. The set 1 scan, which has been described previously [20], included population-based cases and age-matched controls from the 3 population-based centers: Seattle, Toronto, and Australia. Cases were enriched genetically by oversampling those with a young age of onset or positive family history. Controls were matched to cases on age and sex. All cases and controls were self-reported as white, which was confirmed with genotype data. The set 2 scan included population-based cases and matched controls from all 6 colon CFR centers including the Mayo Clinic, Hawaii Cancer Registry, University of Southern California, Fred Hutchinson Cancer Research Center, Ontario Cancer Care, and University of Melbourne. As with set 1, cases were enriched genetically by oversampling persons with a young age at onset or positive family history. Controls were same-generation family controls.

## *Darmkrebs: Chancen der Verhütung durch Screening (DACHS)*

This study was initiated as a large population-based, case-control study in 2003 in the Rhine-Neckar-Odenwald region (southwest region of Germany) to assess the potential of endoscopic screening for reduction of colorectal cancer risk and to investigate etiologic determinants of disease, particularly lifestyle/environmental factors and genetic factors [2,3]. Cases with a first diagnosis of invasive colorectal cancer (International Classification of Diseases 10 codes C18-C20) who were at least 30 years of age (no upper age limit), German speaking, a resident in the study region, and mentally and physically able to participate in a 1-hour interview were recruited by their treating physicians either in the hospital a few days after surgery or by mail after discharge from the hospital. Cases were confirmed based on histologic reports and hospital discharge letters after diagnosis of colorectal cancer. All hospitals treating colorectal cancer patients in the study region participated. Based on estimates from population-based cancer registries, more than 50% of all potentially eligible patients with incident colorectal cancer in the study region were included. Community-based controls were selected randomly from population registries, using frequency matching with respect to age (5-year groups), sex, and county of residence. Controls with a history of colorectal cancer were excluded. Controls were contacted by mail and follow-up telephone calls. The participation rate was 51%. During an in-person interview, data were collected on demographics, medical history, family history of colorectal cancer, and various lifestyle factors, as were blood and mouthwash samples. The set 1 scan consisted of a subset of participants recruited up until 2007, and samples were frequency matched on age and sex. The set 2 scan consisted of additional subjects who were recruited until 2010 as part of this ongoing study.

## *Diet, Activity, and Lifestyle Study (DALS)*

DALS (Diet, Activity, and Lifestyle Study) was a population-based, case-control study of colon cancer. Participants were recruited between 1991 and 1994 from 3 locations: the Kaiser Permanente Medical Care Program of Northern California, an 8-county area in Utah, and the metropolitan Twin Cities area of Minnesota [4]. Eligibility criteria for cases included age at diagnosis between 30 and 79 years; diagnosis with first primary colon cancer (International Classification of Diseases for Oncology second edition codes 18.0 and 18.2–18.9) between October 1, 1991, and September 30, 1994; English speaking; and competency to complete the interview. Individuals with cancer of the rectosigmoid junction or rectum were excluded, as were those with a pathology report noting familial adenomatous polyposis, Crohn's disease, or ulcerative colitis. A rapid-reporting system was used to identify all incident cases of colon cancer, resulting in the majority of cases being interviewed within 4 months of diagnosis. Controls from the Kaiser Permanente Medical Care Program were selected randomly from membership lists. In Utah, controls younger than 65 years of age were selected randomly through random-digit dialing and driver's license lists. Controls, 65 years of age and older, were selected randomly from Health Care Financing Administration lists. In Minnesota, controls were identified from Minnesota driver's licenses or state identification lists. Controls were matched to cases by 5-year age groups and sex. The set I scan consisted of a subset of the study designed earlier, from Utah, Minnesota, and the Kaiser Permanente Medical Care Program, and was restricted to subjects who self-reported as white non-Hispanic. The set 2 scan consisted of subjects from Utah and Minnesota who were not genotyped in set 1. Set 2 was restricted to subjects who self-reported as white non-Hispanic and those who had appropriate consent to post data to dbGaP.

## *Hawaii Colorectal Cancer Studies 2 and 3 (COLO2&3)*

Patients with colorectal cancer were identified through the rapid reporting system of the Hawaii Surveillance, Epidemiology and End Results registry and consisted of all Japanese, Caucasian, and native Hawaiian residents of Oahu who were newly diagnosed with an adenocarcinoma of the colon or rectum between January 1994 and August 1998 [5]. Control subjects were selected from participants in an ongoing population-based health survey conducted by the Hawaii State Department of Health and from Health Care Financing Administration participants. Controls were matched to cases by sex, ethnicity, and age (within 2 years). Personal interviews were obtained from 768 matched pairs, with a participation rate of 58.2% for cases and 53.2% for controls. A questionnaire, administered during an in-person interview, included questions about demographics, lifetime history of tobacco, alcohol use, aspirin use, physical activity, personal medical history, family history of colorectal cancer, height and weight, diet (Food Frequency Questionnaire), and postmenopausal hormone use. A blood sample was obtained from 548 (71%) interviewed cases and 662 (86%) interviewed controls. Surveillance, Epidemiology and End Results staging information was extracted from the Hawaii Tumor Registry. In GECCO, self-reported Caucasian subjects with DNA, as well as clinical and epidemiologic data were selected for genotyping.

## *Health Professionals Follow-up Study (HPFS)*

The HPFS (Health Professionals Follow-up Study) is a parallel prospective study to the NHS (Nurses' Health Study) [6].The HPFS cohort comprised 51,529 men who, in 1986, responded to a mailed questionnaire. The participants were US male dentists, optometrists, osteopaths, podiatrists, pharmacists, and veterinarians born between 1910 and 1946. Participants provided information on health-related exposures, including current and past smoking history, age, weight, height, diet, physical activity, aspirin use, and family history of colorectal cancer. Colorectal cancer and other outcomes were reported by participants or next-of-kin and were followed up through review of the medical and pathology record by physicians. Overall, more than 97% of self-reported colorectal cancers were confirmed by medical record review. Information was abstracted on histology and primary location. Incident cases were defined as those occurring after the subject provided the blood sample. Prevalent cases were defined as those occurring after enrollment in the study but before the subject provided the blood sample. Follow-up evaluation has been excellent, with 94% of the men responding to date. Colorectal cancer cases were ascertained through January 1, 2008. In 1993–1995, 18,825 men in the HPFS mailed blood samples by overnight courier, which were aliquoted into buffy coat and stored in liquid nitrogen. Between 2001 and 2004, 13,956 men in the HPFS who had not provided a blood sample previously mailed in a swish-and-spit sample of buccal cells. Incident cases were defined as those occurring after the subject provided a blood or buccal sample. Prevalent cases were defined as those occurring after enrollment in the study in 1986, but before the subject provided either a blood or buccal sample; these prevalent cases were excluded from our analyses. After excluding participants with histories of cancer (except non-melanoma skin cancer), ulcerative colitis, or familial polyposis, 2 case-control sets were constructed from which DNA was isolated from either buffy coat or buccal cells for genotyping, as follows: (1) a case-control set with cases of colorectal cancer matched to randomly selected controls who provided a blood sample and were free of colorectal cancer at the same time the colorectal cancer was diagnosed in the cases; and (2) a case-control set with cases of colorectal cancer matched to randomly selected controls who provided a buccal sample and were free of colorectal cancer at the same time the colorectal cancer was diagnosed in the case. For both case-control sets, matching criteria included year of birth (within 1 year) and month/year of blood or buccal cell sampling (within 6 months). Cases were pair-matched 1:1, 1:2, or 1:3 with control participant(s).

## *Multiethnic Cohort study (MEC)*

The MEC (Multiethnic Cohort) was initiated in 1993 to investigate the impact of dietary and environmental factors on major chronic diseases, particularly cancer, in ethnically diverse populations in Hawaii and California [7]. The study recruited 96,810 men and 118,441 women aged 45–75 years between 1993 and 1996. Incident colorectal cancer cases occurring since January 1995 and controls were contacted for blood or saliva samples. The median interval between diagnosis and blood draw was 14 months (interquartile range, 10–19 mo) among cases and the participation rate was 74%. A sample of cohort participants was selected randomly to serve as controls at the onset of the nested case-control study (participation rate, 66%). The selection was stratified by sex, age, and race/ethnicity. Colorectal cancer cases were identified through the Rapid Reporting System of the Hawaii Tumor Registry and through quarterly linkage to the Los Angeles County Cancer Surveillance Program. Both registries are members of Surveillance, Epidemiology and End Results. In GECCO, self-reported white subjects from the nested case-control study described earlier with DNA and clinical and epidemiologic data were selected for genotyping.

*Nurses' Health Study (NHS)*

The NHS cohort began in 1976 when 121,700 married female registered nurses age 30–55 years returned the initial questionnaire that ascertained a variety of important health-related exposures [8].Since 1976, follow-up questionnaires have been mailed every 2 years. Colorectal cancer and other outcomes were reported by participants or next-of-kin and followed up through review of the medical and pathology record by physicians. Overall, more than 97% of self-reported colorectal cancers were confirmed by medical-record review. Information was abstracted on histology and primary location. The rate of follow-up evaluation has been high: as a proportion of the total possible follow-up time, follow-up evaluation has been more than 92%. Colorectal cancer cases were ascertained through June 1, 2008. In 1989–1990, 32,826 women in NHS I mailed blood samples by overnight courier, which were aliquoted into buffy coat and stored in liquid nitrogen. In 2001–2004, 29,684 women in NHS I who did not previously provide a blood sample mailed a swish-and-spit sample of buccal cells. Incident cases were defined as those occurring after the subject provided a blood or buccal sample. Prevalent cases were defined as those occurring after enrollment in the study in 1976 but before the subject provided either a blood or buccal sample; these prevalent cases were excluded from our analyses. After excluding participants with histories of cancer (except non-melanoma skin cancer), ulcerative colitis, or familial polyposis, 2 case-control sets were constructed from which DNA was isolated from either buffy coat or buccal cells for genotyping: (1) a case-control set with cases of colorectal cancer matched to randomly selected controls who provided a blood sample and were free of colorectal cancer at the same time the colorectal cancer was diagnosed in the case; and (2) a case-control set with cases of colorectal cancer matched to randomly selected controls who provided a buccal sample and were free of colorectal cancer at the same time the colorectal cancer was diagnosed in the cases. For both case-control sets, matching criteria included year of birth (within 1 year) and month/year of blood or buccal cell sampling (within 6 months). Cases were pair matched 1:1, 1:2, or 1:3 with control participant(s).

## *Ontario Familial Colorectal Cancer Registry (OFCCR)*

In GECCO, a subset of cases and controls included in the Assessment of Risk in Colorectal Tumours in Canada (ARCTIC), and initially recruited for the OFCCR (Ontario Registry for Studies of Familial Colorectal Cancer) was used. Both the case-control study [9] and the OFCCR [10] have been described in detail previously, as have the GWAS results [11]. In brief, cases were confirmed incident colorectal cancer cases if they were ages 20 to 74 years, residents of Ontario, identified through comprehensive registry, and diagnosed between July 1997 and June 2000. Population-based controls were selected randomly among Ontario residents (random-digit dialing and listing of all Ontario residents), and matched by sex and 5-year age groups. A total of 1236 colorectal cancer cases and 1223 controls were genotyped successfully on at least one of the following: Illumina 1536 GoldenGate assay (Illumina, Inc, San Diego, CA), the Affymetrix GeneChip Human Mapping 100K and 500K Array Set (Affymetrix, Inc, Santa Clara, CA), or a 10K nonsynonymous SNP chip. Analysis was based on a set of unrelated subjects who were non-Hispanic, white by self-report, or by investigation of genetic ancestry. We further excluded subjects if there was a sample mix-up, if they were missing epidemiologic questionnaire data, if they were cases with a tumor in the appendix, or if they overlapped with the CCFR. In addition, only samples genotyped on the Affymetrix GeneChip 500K Array were used to avoid coverage issues in imputation. It should be noted that some of the OFCCR sample was included in the CCFR group.

*Physicians' Health Study (PHS)*

The PHS (Physicians' Health Study) was established as a randomized, double-blind, placebo-controlled trial of aspirin and β-carotene among 22,071 healthy US male physicians, between 40 and 84 years of age, in 1982 [12,13].Participants completed 2 mailed questionnaires before being assigned randomly, additional questionnaires at 6 and 12 months, and questionnaires annually thereafter. In addition, participants were sent postcards at 6 months to ascertain status. From August 1982 to December 1984, there were 14,916 baseline blood samples collected from the physicians during the run-in phase before randomization. When participants reported a diagnosis of cancer, medical records and pathology reports were reviewed by study physicians who were blinded to exposure data. Among those who provided baseline blood samples, colorectal cases were ascertained through March 31, 2008, and controls were matched on age (within 1 year for younger participants, up to 5 years for older participants) and smoking status (never, past, current). Cases were pair-matched 1:1, 1:2, or 1:3 with control participant(s). Because of DNA availability, samples were genotyped in 2 batches on the same platform at the same genotyping center at different time points.

## *Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO)*

The PLCO (Prostate, Lung, Colorectal Cancer, and Ovarian Cancer Screening Trial) enrolled 154,934 participants (men and women, aged between 55 and 74 y) at 10 centers into a large, randomized, 2-arm trial to determine the effectiveness of screening to reduce cancer mortality. Sequential blood samples were collected from participants assigned to the screening arm. Participation was 93% at the baseline blood draw. In the observational (control) arm, buccal cells were collected via mail using the swish-and-spit protocol; the participation rate was 65%. Details of this study have been described previously [14,15] and are available online ([http://dcp.cancer.gov/plco](http://www.sciencedirect.com/science?_ob=RedirectURL&_method=externObjLink&_locator=url&_issn=00165085&_origin=article&_zone=art_page&_plusSign=%2B&_targetURL=http%253A%252F%252Fdcp.cancer.gov%252Fplco)).

The set 1 scan included a subset of 577 colon cancer cases self-reported as being non-Hispanic white with available DNA samples, questionnaire data, and appropriate consent for ancillary epidemiologic studies. Cases were excluded if they had a history of inflammatory bowel disease, polyps, polyposis syndrome, or cancer (excluding basal or squamous cell skin cancer). Controls originated from the Cancer Genetic Markers of Susceptibility prostate cancer scan [16,17] (all male) and the GWAS of Lung Cancer and Smoking [18] (enriched for smokers), along with an additional 92 non-Hispanic white female controls. Set 1 samples have been excluded from this analysis. For the set 2 scan, cases were individuals with colorectal cancer from both arms of the trial who were not already included in set 1. Samples were excluded if participants did not sign appropriate consent forms, if DNA was unavailable, if baseline questionnaire data with follow-up evaluation were unavailable, if they had a history of colon cancer before the trial, if they had a rare cancer, if they were already in a colon GWAS, or if they were a control in the prostate or lung populations. Controls were frequency-matched 1:1 to cases without replacement, and cases were not eligible to be controls. Matching criteria were age at enrollment (2-year blocks), enrollment date (2-year blocks), sex, race/ethnicity, trial arm, and study year of diagnosis (ie, controls must be cancer free into the case's year of diagnosis).

## *Postmenopausal Hormones Supplementary Study to the CCFR (PMH-CCFR)*

Eligible case patients included all female residents, ages 50–74 years, residing in the 13 counties in Washington State, reporting to the Cancer Surveillance, Epidemiology and End Results program, who were newly diagnosed with invasive colorectal adenocarcinoma (ICD-O C18.0, C18.2–C18.9, C19.9, C20.0–C20.9) between October 1998 and February 2002 [19]. Eligibility for all individuals was limited to those who were English speaking with available telephone numbers, through which they could be contacted. On average, cases were identified within 4 months of diagnosis. The overall response proportion of eligible cases identified was 73%. Community-based controls were selected randomly according to age distribution (in 5-year age intervals) of the eligible cases by using lists of licensed drivers from the Washington State Department of Licensing for individuals, ages 50–64 years, and rosters from the Health Care Financing Administration (now the Centers for Medicare and Medicaid), for individuals older than age 64. The overall response proportion of eligible controls was 66%. In GECCO, samples with sufficient DNA extracted from blood were genotyped. Only participants who were not part of the CCFR Seattle site were included in the sample set.

## *VITamins And Lifestyle (VITAL)*

The VITAL (VITamins And Lifestyle) cohort comprised 77,721 Washington State men and women aged 50–76 years, recruited from 2000 to 2002, to investigate the association of supplement use and lifestyle factors with cancer risk. Subjects were recruited by mail, from October 2000 to December 2002, using names purchased from a commercial mailing list. All subjects completed a 24-page questionnaire and buccal-cell specimens for DNA were self-collected by 70% of the participants. Subjects were followed for cancer by linkage to the western Washington Surveillance, Epidemiology and End Results (SEER) cancer registry and were censored when they moved out of the area covered by the registry or at time of death. Details of this study have been described previously [20]. In GECCO, a nested case-control set was genotyped. Samples included colorectal cancer cases with DNA, excluding subjects with colorectal cancer before baseline, in situ cases, (large cell) neuroendocrine carcinoma, squamous cell carcinoma, carcinoid tumor, Goblet-cell carcinoid, and any type of lymphoma, including non-Hodgkin, Mantle cell, large B-cell, or follicular lymphoma. Controls were matched on age at enrollment (within 1 year), enrollment date (within 1 year), sex, and race/ethnicity. One control was selected randomly per case among all controls who matched according to the 4 factors described earlier and for whom the control follow-up time was greater than the follow-up time of the case until diagnosis.

## *Women's Health Initiative (WHI)*

The WHI (Women's Health Initiative) is a long-term health study of 161,808 post-menopausal women aged 50–79 years at 40 clinical centers throughout the United States. WHI comprised a clinical trial arm, an observational study (OS) arm, and several extension studies. The details of WHI have been described previously [21,22]and are available online ([https://cleo.whi.org/SitePages/Home.aspx](http://www.sciencedirect.com/science?_ob=RedirectURL&_method=externObjLink&_locator=https&_issn=00165085&_origin=article&_zone=art_page&_plusSign=%2B&_targetURL=https%253A%252F%252Fcleo.whi.org%252FSitePages%252FHome.aspx)). In GECCO, set 1 cases were selected from the September 12, 2005, database and comprised centrally adjudicated colon cancer cases from the OS arm who self-reported as white. Controls were first selected among controls previously genotyped as part of a hip fracture GWAS conducted within the WHI OS arm and matched to cases on age (within 3 years), enrollment date (within 365 days), hysterectomy status, and prevalent conditions at baseline. For 37 cases, there was no control match in the hip fracture GWAS. For these participants, we identified a matched control in the WHI OS arm based on the same criteria. In the set 2 scan, cases were selected from the August 2009 database and comprised centrally adjudicated colon and colorectal cancer cases from the OS and clinical trial arms who were not genotyped in set 1. In addition, case and control participants were subject to the following exclusion criteria: a prior history of colorectal cancer at baseline, institutional review board approval not available for data submission into dbGaP, and insufficient DNA available. Matching criteria included age (within 3 years), race/ethnicity, WHI date (within 3 years), WHI Calcium and Vitamin D study date (within 3 years), and randomization arms (OS flag, hormone therapy assignments, dietary modification assignments, calcium/vitamin D assignments). In addition, study participants were matched on randomization centers, with analytic adjustment made for the 4 regions of randomization centers. Each case was matched with 1 control (1:1) who met the matching criteria exactly (i.e., cases and controls were matched on trial assignments and on randomization within these trials; in the rare event that a case is in 2 trials, she was matched to a control in the same two trials). Control selection was performed in a time-forward manner, selecting one control for each case first from the risk set at the time of the case's event. The matching algorithm was allowed to select the closest match based on a criterion to minimize an overall distance measure [23]. Each matching factor was given the same weight. Additional available controls who were genotyped as part of the hip fracture GWAS were included to improve power.

SUPPLEMENTAL REFERENCES

1. Newcomb PA, Baron J, Cotterchio M, Gallinger S, Grove J, [Haile R](http://www.ncbi.nlm.nih.gov/pubmed?term=Haile%20R%5BAuthor%5D&cauthor=true&cauthor_uid=17982118), et al. Colon Cancer Family

Registry: an international resource for studies of the genetic epidemiology of colon cancer. Cancer

Epidemiol Biomarkers 2007; Prev 16: 2331-43.

2. Brenner H, Chang-Claude J, Seiler CM, Rickert A, Hoffmeister M. Protection from colorectal

cancer after colonoscopy: a population-based, case-control study. Ann Intern Med 2011; 154: 22-30.

3. Lilla C, Verla-Tebit E, Risch A, [Jäger B](http://www.ncbi.nlm.nih.gov/pubmed?term=J%C3%A4ger%20B%5BAuthor%5D&cauthor=true&cauthor_uid=16434594), [Hoffmeister M](http://www.ncbi.nlm.nih.gov/pubmed?term=Hoffmeister%20M%5BAuthor%5D&cauthor=true&cauthor_uid=16434594), [Brenner H](http://www.ncbi.nlm.nih.gov/pubmed?term=Brenner%20H%5BAuthor%5D&cauthor=true&cauthor_uid=16434594), et al. Effect of NAT1 and NAT2

Genetic polymorphisms on colorectal cancer risk associated with exposure to tobacco smoke and

meat consumption. Cancer Epidemiol Biomarkers Prev 2006; 15: 99-107.

4. Slattery ML, Potter J, Caan B, Edwards S, Coates A, [Ma KN](http://www.ncbi.nlm.nih.gov/pubmed?term=Ma%20KN%5BAuthor%5D&cauthor=true&cauthor_uid=8988044) et al. Energy balance and colon

cancer- beyond physical activity. Cancer Res 1997; 57: 75-80.

5. Le Marchand L, Hankin JH, Wilkens LR, Pierce LM, Franke A, [Kolonel LN](http://www.ncbi.nlm.nih.gov/pubmed?term=Kolonel%20LN%5BAuthor%5D&cauthor=true&cauthor_uid=11751443), et al. Combined effects

of well done red meat, smoking, and rapid N-acetyltransferase 2 and CYP1A2 phenotypes in increasing

colorectal cancer risk. Cancer Epidemiol Biomarkers Prev 2001; 10: 1259-66.

6. Rimm EB, Stampfer MJ, Colditz GA, Chute CG, Litin LB, [Willett WC](http://www.ncbi.nlm.nih.gov/pubmed?term=Willett%20WC%5BAuthor%5D&cauthor=true&cauthor_uid=2090285). Validity of self-reported

waist and hip circumferences in men and women. Epidemiology 1990; 1: 466-73.

7. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, [Pike MC](http://www.ncbi.nlm.nih.gov/pubmed?term=Pike%20MC%5BAuthor%5D&cauthor=true&cauthor_uid=10695593), et al. A multiethnic

cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol 2000; 151: 346-57.

8. Belanger CF, Hennekens CH, Rosner B, Speizer FE. The nurses' health study. Am J Nurs 1978; 78:

1039-40.

9. Cotterchio M, Manno M, Klar N, McLaughlin J, Gallinger S. Colorectal screening is associated

with reduced colorectal cancer risk: a case-control study within the population-based Ontario

Familial Colorectal Cancer Registry. Cancer Causes Control 2005; 16: 865-75.

10.Cotterchio M, McKeown-Eyssen G, Sutherland H, Buchan G, Aronson M, [Easson AM](http://www.ncbi.nlm.nih.gov/pubmed?term=Easson%20AM%5BAuthor%5D&cauthor=true&cauthor_uid=11007659), et al.

Ontario familial colon cancer registry: methods and first-year response rates. Chronic Dis Can 2000;

21: 81-6.

11.Zanke BW, Greenwood CM, Rangrej J, Kustra R, Tenesa A,  [Farrington SM](http://www.ncbi.nlm.nih.gov/pubmed?term=Farrington%20SM%5BAuthor%5D&cauthor=true&cauthor_uid=17618283), et al. Genome-wide

Association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat Genet

2007; 39: 989-94.

12.Hennekens CH, Eberlein K. A randomized trial of aspirin and beta-carotene among U.S.

physicians. Prev Med 1985; 14: 165-8.

13.Christen WG, Gaziano JM, Hennekens CH. Design of Physicians' Health Study II--a randomized

trial of beta-carotene, vitamins E and C, and multivitamins, in prevention of cancer, cardiovascular

disease, and eye disease, and review of results of completed trials. Ann Epidemiol 2000; 10: 125-34.

14.Gohagan JK, Prorok PC, Hayes RB, Kramer BS et al. The Prostate, Lung, Colorectal

and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization,

and status. Control Clin Trials 2000; 21: 251S-72S.

15.Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, [Crawford ED](http://www.ncbi.nlm.nih.gov.ezp-prod1.hul.harvard.edu/pubmed?term=Crawford%20ED%5BAuthor%5D&cauthor=true&cauthor_uid=11189684) et al. Design of the

Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. Control Clin Trials 2000; 21:

273S-309S.

16.National Cancer Institute [Internet]. Cancer Genetic Markers of Susceptibility (CGEMS) data website.

Available from: http://cgems.cancer.gov/data\_access.html. CGEMS Data accessed October 5, 2009.

17.Yeager M, Chatterjee N, Ciampa J, Jacobs KB, Gonzalez-Bosquet J, [Hayes RB](http://www.ncbi.nlm.nih.gov/pubmed?term=Hayes%20RB%5BAuthor%5D&cauthor=true&cauthor_uid=19767755), et al. Identification of a

new prostate cancer susceptibility locus on chromosome 8q24. Nat Genet 2009; 41: 1055-7.

18.Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM,  [Rotunno M](http://www.ncbi.nlm.nih.gov/pubmed?term=Rotunno%20M%5BAuthor%5D&cauthor=true&cauthor_uid=19836008), et al. A genome-wide

association study of lung cancer identifies a region of chromosome 5p15 associated with risk for

adenocarcinoma. Am J Hum Genet 2009; 85: 679-91.

19.Newcomb PA, Zheng Y, Chia VM, Morimoto LM, Doria-Rose VP, [Templeton A](http://www.ncbi.nlm.nih.gov/pubmed?term=Templeton%20A%5BAuthor%5D&cauthor=true&cauthor_uid=17671225), et al. Estrogen

plus progestin use, microsatellite instability, and the risk of colorectal cancer in women. Cancer Res

2007; 67: 7534-9.

20.White E, Patterson RE, Kristal AR, Thornquist M, King I, [Shattuck AL](http://www.ncbi.nlm.nih.gov/pubmed?term=Shattuck%20AL%5BAuthor%5D&cauthor=true&cauthor_uid=14693663), et al. VITamins And Lifestyle

cohort study: study design and characteristics of supplement users. Am J Epidemiol 2004; 159: 83-93.

21.Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, [Allen C](http://www.ncbi.nlm.nih.gov/pubmed?term=Allen%20C%5BAuthor%5D&cauthor=true&cauthor_uid=14575939), et al. The Women's Health Initiative

recruitment methods and results. Ann Epidemiol 2003; 13: S18-77.

22.The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and

observational study. Control Clin Trials 1998; 19: 61-109.

23. Bergstralh EJ, Kosanke JL. Computerized matching of cases to controls. Rochester (MN): Department

of Health Sciences Research, Mayo Clinic; 1995. Technical Report No.: 56.