

Canadian Partnership for Tomorrow Project



BC Generations Project Protocol

REB # H08-01354

Document Change Log

Version	Date	Description	Done by
1.7	Sep 24, 2009	Vanguard protocol	R.Gallagher
1.8	Jan 14, 2010	Phase 2 protocol	R. Gallagher
2.0	July 19, 2010	National on-line Q	M.Borugian
2.1	Jan 14, 2011	Chg age range to 35-69	M.Borugian
2.2	April 4, 2011	Remove recruitment end date of 2012	E.Rousseau

Conducted by the BC Cancer Agency
in association with
The Canadian Partnership Against Cancer
and
The BC Cancer Foundation

BC Generations Project Protocol

Summary of April 4, 2011 Modifications

While development of the on-line questionnaire is continuing, the project will continue to enrol participants and collect paper based questionnaires. When the on-line questionnaire is ready, we will offer that as an additional method for participation. We are offering participation at an assessment centre at the BCCRC one day per week, and hope to operate assessment centres in major cities throughout British Columbia.

Summary of July 19, 2010 Modifications

In response to participant feedback, and to mitigate financial issues arising from the economic downturn, the Principal Investigators from the 5 participating regions (BC, Alberta, Ontario, Quebec and the 4 Atlantic provinces) have agreed upon a single 'core' national questionnaire and have further agreed to offer the option to complete the questionnaire on-line. Despite 'common wisdom' that the target age group is not willing to use computer technology, our experience has been that over 90% of participants provide email addresses and dozens of feedback forms have chastised us for using too much paper. To respond to participant feedback, and to manage costs, we therefore have designed an on-line system to handle participant registration, participant informed consent, and completion of the new national questionnaire. The detailed explanation of the process, along with information on the critical issues of identifying information and protection of privacy utilizing appropriate security techniques can be found in the Methods section, below.

Introduction to 2010 Revised Protocol

This document will outline a revised protocol for the BC Generations Project based on the experience gained in recruiting the initial 4000 of 40,000 participants using the BC Generations Project Protocol Version 1.7 (September 24, 2009) as approved by the UBC-BCCA Research Ethics Board. This original protocol outlined B.C.'s contribution to a Pan Canadian effort to recruit 300,000 Canadians age 35-69 into a prospective health cohort which would serve as a research platform for investigation of disease causation and biomarker discovery for scientists across Canada.

The September 24, 2009 protocol envisaged use of clinical assessment centres travelling throughout the province and collecting questionnaire data, as well as blood and urine specimens and conducting physical measurements on participants. The first 4000 participants were collected through the first assessment centre in Vancouver at the Gordon and Leslie Diamond Health Care Centre. While this procedure proved popular with participants and was very effective, it has also proven to be very expensive.

When the decision was made in early 2008 to use assessment centres, it was anticipated that substantially more funding than that available from the Canadian Partnership Against

Cancer could be raised during 2009 to provide for the completion of the recruitment phase of this project. Unfortunately the economic downturn which began in the U.S. in September 2008 subsequently spread throughout Western countries in 2009, including Canada, resulting in a substantial shortfall for continuing the project on the former basis. It is emphasized that this problem has affected each of the provincial/regional cohorts; Alberta, Ontario, Quebec and the Atlantic provinces, in a similar fashion in the Canadian Partnership for Tomorrow. Faced with the choice of reducing recruitment goals or revising the protocol, all the cohorts comprising the Canadian Partnership for Tomorrow have elected to collect the same number of participants as envisaged in the initial protocol and within the same time frame but without necessarily using assessment centres and with a somewhat ‘thinner’ dataset at initial enrolment.

This does not mean that the BC Generations Project and the other provincial/regional cohorts have decided to abandon any of the original project goals—simply that collection of specimens (blood and urine) and physical measurements from some participants will occur at a time subsequent to the initial recruitment phase, as more funds become available.

The revised protocol presented here represents our ‘go forward’ methods for the time period beginning February 2010. As noted above, the nature of the data, specimens and measurements to be collected remain the same as in the September 24, 2009 protocol. In addition, the recruitment techniques used with the first 4000 participants remain the same in the new approach—the changes are in the method used to collect the data.

The major changes are:

- 1) The BC Generations Project will discontinue use of clinical assessment centres to collect data, blood, urine, and physical measurements. This approach has proven to be too expensive to sustain at this time. Instead, participants will complete the informed consent and the enrolment questionnaires at home either on paper or (later as we acquire the capability) online at their preference. The informed consent and paper questionnaire will be mailed back to the BC Generations Project. A laboratory requisition will be included for as many of the participants that our budget will accommodate, asking them to have blood drawn and a urine specimen taken at a LifeLabs laboratory close to their home or workplace. LifeLabs will ship the specimens to our newly established processing lab at the BC Cancer Research Centre for processing and storage. We are currently seeking funds for a mobile operation to collect physical measurement data on participants, but if no funds are obtained, assessment of physical measurements will be deferred until later follow-up contact with participants after recruitment is complete. The project protocol asks for permission for re-contact from each participant and, to date, has not had any refusals. We will also begin to contact some large and small companies, firms, and perhaps unions in British Columbia to seek recruits from within the companies. This has proved successful in pilot studies conducted in other cohorts in the past.

- 2) With the closing of our assessment centre at the Gordon and Leslie Diamond Healthcare Centre on February 11, 2010, blood processing will be carried out at the BC Cancer Research Centre. Techniques for processing, preserving, and storing the specimens in a de-identified form will remain the same.

All other aspects of the study remain essentially unchanged.

Rationale

The Importance of Chronic Disease Prevention

Despite the major progress in cancer detection and treatment made in the last 10 years, the most cost-effective and sustainable way of controlling cancer remains prevention. There is little disagreement that preventing cancer is the ultimate goal of cancer control. However to prevent a disease it is necessary to identify the cause, and for many cancers, the cause is either unknown or controversial. What is known is that chronic diseases such as cancer are caused by a combination of lifestyle factors, exposure to environmental agents, and individual genetic make up. In cases where the environmental factor is known (e.g. use of tobacco products), public health interventions and education programs are succeeding in reducing the level of exposure (smoking) and hence the incidence of lung cancer. It is estimated that the complete removal of tobacco products would reduce overall cancer incidence by as much as 30% (Harvard Center for Cancer Prevention, 1996), and that most skin cancers could be prevented by reducing overexposure to ultraviolet radiation from sunlight and tanning beds (IARC 1992; IARC 2001). However, the effect of even a well known carcinogen, such as sunlight, can differ by up to 100-fold among populations of different genetic make-up (Rees, 2004). Thus, for cancer prevention to have a significant impact at the population level, an understanding of the environmental risk factors involved, and the interplay between these factors and human genetic make-up, is needed.

Approach: The Need for a ‘Big Science’ in Disease Prevention Research

Until recently, risk factors for cancer and most other chronic diseases have been identified primarily using case-control studies, in which the characteristics of groups of patients (cases) were compared to those reported by age and gender matched groups of people who had not developed cancer (controls). In such studies, subjects are typically asked to complete detailed questionnaires about their lifestyle, environmental exposures, diet, socio-economic status, medical history, etc. in some pre-defined period prior to cancer diagnosis. Investigators in various parts of the world, generally acting alone, or in small local groups, have designed questionnaires and conducted studies independently. While this has resulted in a substantial body of literature, it has produced (with some notable exceptions) a host of contradictory findings due to differences in study design, sampling procedures, study power, information format, participant recall, and other issues (Smith and Ibrahim, 2001; LeFanu 1999). Furthermore, the lack of uniform data

acquisition instruments and study methods to collect etiologic information has made knowledge synthesis through conduct and analysis of ‘pooled’ datasets problematic. For the past 10 years, investigators have sought to obtain blood samples for extraction of DNA in order to look for genetic variations that may help predict disease risk. Again, much of this activity has been conducted independently and has resulted in a huge literature of positive and negative findings for individual genes. Many of what were thought to be promising findings have proven not to be reproducible (Terry and Goodman, 2006).

Although the need for “big science” in complex scientific areas such as nuclear physics became evident over 60 years ago during the Manhattan Project (Rhodes, 1986), its relevance to health sciences, and particularly to cancer and chronic disease prevention, has become apparent only recently (Hoover, 2007). However, the sequencing of the human genome (International Human Genome Sequencing Consortium, 2001; Venter *et al*, 2001) established the need for major collaborative efforts to solve health science problems. The realization that there are more than 30,000 genes, with potentially millions of common variants, drove home the need for change in the cancer research paradigm; moving away from projects driven by single isolated investigators or small teams, and moving towards large interdisciplinary consortia of scientists working toward a common goal.

Recently, epidemiologists, biostatisticians, and population geneticists (the ‘basic scientists’ of cancer and chronic disease prevention) have also moved to the formation of large coordinated interdisciplinary groups to study environmental and lifestyle risk factors for disease. The InterLymph Consortium (<http://epi.grants.cancer.gov/InterLymph/>) established by the US National Cancer Institute has brought together principal investigators from more than 20 studies all over the world to coordinate the investigation of environmental and genetic risk factors for non-Hodgkin lymphoma (NHL), an entity which, in actuality, comprises some 35 different diseases. None of the NHL studies alone would have been able to accrue sufficient numbers of subjects to study the many disease variants separately but, through collaboration and data pooling, understanding of the environmental and genetic factors involved in these diseases is evolving rapidly (Rothman *et al*, 2006; Morton *et al*, 2007; Wang *et al*, 2007; Krickler *et al*, 2007). The elaboration of ‘special study groups’ within InterLymph (behaviour and environment, host and genetics, pathology and survival, etc.) has also ensured that data are aggregated uniformly, and that optimal use is made of information available from each study site. Similar international consortia are engaged in the study of the causes of melanoma (Genes Environment and Melanoma Consortium; GEM); brain cancer (Brain Tumour Epidemiology Consortium; BTEC); and other tumours.

In the United States the use of consortia in the study of cancer is receiving a great deal of attention (Potter, 2004; Collins 2004; Hoover, 2007) and efforts are underway to combine existing compatible cohort studies (<http://epi.grants.cancer.gov/Consortia/cohort.html>) in order to study genetic, environmental, and lifestyle findings in disease. In addition, the dramatic (and

continuing) fall in genotyping costs over the past five years has made the ‘full-genome scan’ the method of choice over the ‘candidate gene’ approach in determining the effect of genetic variants on human disease. This in turn has further encouraged the need for the consortia approach in order to ensure the large numbers of subjects needed for hypothesis development datasets and independent confirmation datasets (Hunter *et al*, 2007). A further benefit of these large scale datasets is that they minimize the problem of ‘false positives’ inherent in association-testing using literally hundreds of thousands of genetic variants.

Finally, the knowledge that environmental and lifestyle exposures may give rise to epigenetic changes in the genome, which substantially modify risk for cancer even though they may not change gene sequence, offers a new field of investigation in which collaborative studies (and particularly cohort studies) will be singularly valuable scientific platforms (Jones and Baylin, 2007). The aging process and its associated environmental insults (and perhaps multigenerational processes and exposures) appear to be critically important in determining disease risk ((Mathers, 2007; Weidman et al, 2007; Cobiak, 2007). In fact, epigenomics might offer a unique opportunity for geneticists and environmental and lifestyle disease epidemiologists to collaborate more closely.

The establishment of a Canadian cohort (formally called the Canadian Partnership for Tomorrow, CPFT) in response to these global trends in research will ensure Canada has a key scientific platform necessary to take advantage of, and be a full participant in, the emerging scientific revolution in determining the causes of cancer and chronic diseases. The BC component (called BC Generations Project) will be part of this initiative.

With many health cohorts already underway throughout the world, serious questions might be raised as to why Canada needs health cohorts. There is currently a major void in research specifically aimed at exploring modifiable environmental and lifestyle factors using modern methods of exposure measurement. This is a void that Canada can fill. The CPFT will accomplish what no other major international groups are doing: specifically building a prospective cohort platform that will focus primarily on the investigation of environmental and lifestyle risk factors for cancer and their interaction with genetic and epigenetic risk factors. Aspects of our environment and lifestyle, such as water quality, air quality and activity levels, have seldom been thoroughly evaluated in epidemiologic cohorts, because their measurement is problematic. However, a new Canadian Cancer Cohort would be ideally placed to take advantage of the emergence of new technologies, as well as the growing realization amongst epidemiologists that challenging issues can be addressed in innovative ways using trans-disciplinary approaches.

Operational Objectives of the CPFT and BC Generations Project

- 1) To build the Canadian Partnership for Tomorrow cohort comprised of a confederation of provincial cohorts currently under construction, or underway, in Alberta, Atlantic Canada, British Columbia (BC Generations Project), Ontario and Quebec. The current status of each of the provincial initiatives is described in the next section.

- 2) To enrol between 250,000 and 300,000 Canadians in the CPFT over the next five years, beginning with the current provincial initiatives, including the potential for adding other provinces/territories.
- 3) To collect high quality lifestyle and environmental exposure information as well as high quality DNA and other biological samples (e.g. plasma, serum, urine) and physical measurements from the enrolled cohort members.
- 4) To build the facility for use of data and specimens by scientists across Canada, and potentially through collaborations with other cohorts to scientists around the world.
- 5) To make the links necessary to ensure the participation of the CPFT in international collaborative disease studies around the world. Anonymized data on individuals will be shared with other cohorts
- 6) To find sufficient financial backing to ensure good follow-up of the cohort over the 20 years following completion of recruitment and initial data/sample collection.
- 7) To undertake all activities within a clear and transparent legal framework of accountability and with the highest ethical standards.

Objectives of the BC Generations Project

- 1) To recruit a total of 40,000 BC residents age 35-69 years into the BC Generations Project by collecting personal information.
- 2) To acquire, process and store DNA blood fractions and urine, as well as physical measurements from all participants in the BC Generations Project.
- 3) To acquire participant permission for use of the data and specimens for cancer and other chronic disease research.
- 4) To acquire participant permission to access past health history information and to do follow-up through linkage to administrative health records.
- 5) To ensure that data and specimen collection procedures and products are harmonized, under controlled conditions, with those used in other provinces in order to be able to maximize the value of the cohort for the study of disease. This is necessary because many individual cancers are relatively rare, and the power to study diseases increases much more rapidly by combining datasets.

Methods: BC Generations Project

Overview and timeline

The methods will not cover aspects of the BC cohort involving harmonization with the other components of the CPFT. These can be found in the CPFT protocol as approved by the Canadian Partnership Against Cancer.

The aim of BC Generations Project is to recruit 40,000 BC residents age 35-69. Each participant will be asked to give informed consent, and asked to complete an etiologic questionnaire. Participants will also be asked donate a 40 ml blood specimen and a urine specimen either immediately or at a convenient time in the future. In addition some

participants will have several physical parameters (blood pressure, bone density, lung function, body composition, grip strength) measured. All information is collected and used in strict accordance with the the BC *Freedom of Information and Protection of Privacy Act* (FIPPA) and other applicable legislation, and subject to the independent oversight of the Office of the BC Information and Privacy Commissioner.

Information collected will be entered into electronic form as soon as possible after collection to promote data quality. Information will be stored in secure electronic form at the BC Cancer Agency in encrypted and password protected files. Personal identifying information will be stripped off; only the participant's unique study number will be maintained on the file. An electronic key linking participant name, address, and phone number, with the cohort ID number, (same as the unique study number) will be stored in an encrypted form in a separate password protected file on a different server. This key file will be transferred to a separate zip drive, which will be kept in a safe within the BC Cancer Agency with extremely limited access to it.

Specimens will be bar coded with the participant's unique study number only. No personal identifying information will be present on the tubes. The specimen collection data, with participant unique study ID, will be stored on a separate server from the questionnaire data. The same electronic key used with questionnaire data will enable linkage with the participant's identity. The key as noted will be stored under password protection in an encrypted fashion on a third server.

Discussions are underway with Population Data BC to enable information from BC cohort participants to be stored at their secure long-term storage site at the University of British Columbia (www.popdata.bc.ca). This practice will facilitate linkage with other health information within Population Data BC, and will take advantage of the very tight physical and electronic security measures in place there.

Approved researchers will be able to apply to use the anonymous linked data files to investigate the relationship between environmental, lifestyle and susceptibility factors and chronic diseases. A draft data policy for access only to BC Generations Project data has been written. A protocol for health researchers to apply for use of data from all CPFT cohort sites is currently under development by principal investigators of the 5 provincial cohorts under the chairmanship of Dr. Bartha Maria Knoppers of the Université de Montréal.

Timeframe: British Columbia

Period 1 - 14th January, 2008 to 31st March 2008

- acquire needed equipment
- start to develop harmonized best practice protocols and operating procedures for IT, acquiring and storing core data, and acquiring and storing specimens in collaboration with PIs of the other provincial cohorts and content experts.

Period 2 - 1st April 2008 to 31st May 2009

- develop, in collaboration with other cohort PIs, the framework for permitting use of the data and specimens.
- collect specimens and datasets from the vanguard group of approximately 300 cohort participants enrolled through I-HELP (REB # H07-00740) at the first cohort Assessment Centre in the Greater Vancouver Regional District (GVRD).

Period 3 - 1st June 2009 to 31st February 2010

- recruit a further 4,000 cohort participants with bio-specimens datasets in the GVRD by bringing participants into a clinical assessment centre at the Gordon and Leslie Diamond Health Care Centre in Vancouver as described in the Protocol 1.7 September 24, 2009.

Period 4 – 1st March 2010 to 31st December, 2010

- Began new phase of recruitment utilizing paper questionnaires and blood collection at local phlebotomy clinics. Continued to accrue participants to a small assessment centre at the BC Cancer Research Centre.

Period 5 – 1st January 2011 to present

- Continued to recruit participants using paper questionnaires and the BCCRC assessment centre. Recruited participants into a clinical assessment centre in Kelowna for 5 weeks in January and February.

Future

- Develop plans to open additional clinical assessment centres in locations around British Columbia. Current plans include assessment centres in Prince George, Abbotsford and Victoria within the next year. Continue to explore and develop the online questionnaire procedure with hopes of implementation during fiscal 2011-2012.

Recruitment

Direct mailing: In our initial recruitment (June 2009-February 2010), letters addressed to individuals resident within the 16 km diameter catchment area of our assessment centre at the Gordon and Leslie **Diamond Healthcare Centre** were sent out in batches of 5,000. The letter included a brief description of the project, and an invitation to either call or e-mail the BC Generations Project to join. Participants were able to find out more information either by telephone contact with BC Generations, or by looking up more on our website (www.bcgenerationsproject.ca). Names for direct mailing were provided by Info Canada Ltd directly to PDQ Print Solutions Ltd, who packaged and mailed invitations to potential participants.

In addition to direct mailing, project personnel will approach large companies, public sector bodies such as hospitals, and unions to seek permission to recruit participants from within these organizations using company newsletters etc. Information given to these potential participants will be the same as that sent to participants recruited through direct mailing.

The first 4000 age-eligible participants visited our assessment centre between June 2009 and January 2010 and completed all aspects of enrolment. From February 2010 further recruitment will also be conducted using paper or online questionnaires, with blood and urine specimens collected at LifeLabs laboratories close to the homes or workplaces of the participants. During the period 2010-12 Project will recruit participants from the Vancouver region and other areas of the province in order to ensure the final cohort is as representative as possible of the provincial population.

Project personnel are seeking money from private sources to fund a mobile trailer staffed by an RN and an assistant, which will move through the main population centres of the province to collect physical measurement data on as many participants as possible. If no money is found for this, the project will collect physical measurement data after recruitment is finished, as funds become available.

Random digit dialing:

Telephone calls to residences were made by eNRG Research Group Ltd in Vancouver on a pilot basis in June –July 2009. Costs for this service were very high on a per-recruit basis. Because of this, a decision was made to discontinue this method of recruitment. The principal investigator of the BC Generations Project has requested that the BC Ministry of Health Services send out invitations for participation on behalf of the project, and discussions are underway with the Ministry. However, at the present time no agreement has been reached with the government and the project will proceed in the meantime with the recruitment methods noted above.

.

Invitations

Once participants - located through direct mailing, or through companies hospitals and unions- indicate interest in the project, they are sent a copy of the participant information and consent form. This form describes, in a transparent way, how data and specimens would be taken and used, and the risks and benefits involved. The mailing also includes the location of our website.

BC Generations Project Website

In order to provide as complete information as possible to potential participants, we have constructed a website (www.bcgenerationsproject.ca) which includes full information about the project. Although the website does not currently include interactive functions, and will not allow data entry, access or modification, it contains:

- a) The study description, including the protocol for the Canadian Partnership for Tomorrow and the protocol for BC Generations Project
- b) The letter of invitation
- c) An outline of how people are selected for invitation (also in the BC Generations Project protocol)
- d) Study documents including the Participant Information and Consent form and the Privacy Impact Assessment.
- e) Contact information
- f) Frequently asked questions and answers

The website will be updated frequently in order to keep participants up to date with the progress of BC Generations Project.

Visit of the initial 4000 Participants to the BC Generations Project Assessment Centre

As noted earlier, it was originally anticipated that all 40,000 participants visit a BC Generations Project assessment centre for all aspects of enrolment including questionnaire completion, blood draw, urine donation and physical measurements. This approach would minimize the travel and time commitment a participant must make in order to join the cohort, and ensure completion of all aspects of cohort entry at the same time. As noted earlier, the first assessment centre was located on the 6th floor of the Gordon and Leslie Diamond Health Care Centre in Vancouver.

At entry participants reviewed and signed the electronic Participant Consent Form, and completed each of the following:

- a) The baseline information questionnaire.

- b) Donation of a blood sample (40 ml) to provide the project with DNA, serum, and plasma.
- c) Underwent a set of non-intrusive baseline measurements (height, weight, body fat, respiratory function, blood pressure, grip strength, bone density)
- d) Gave permission for the project to access past and future health information through linking with routinely collected health databases including the BC Cancer Registry and Ministry of Health Services files, for research and statistical purposes only.
- e) Gave permission to store their specimens and information and utilize them (even after death or disablement) anonymously for bona fide research projects.
- f) Gave permission to cohort personnel to re-contact them in the future to request further information or specimens.

While this method has worked extremely well, it has proven to be much more expensive than the BC Generations Project can afford with current funding. To actually recruit the 40,000 participants, we have revised our procedures, as detailed below.

Revised Protocol

Letters (in batches of 5000-10000) will be sent out as described above, inviting participation in the project. The letters describe the project, and invite participants to either phone our call centre for more information or access our website to enrol.

Currently only the paper option exists, but in future participants will be offered their choice of a paper or an on-line enrolment package. The paper package remains as described previously, with the exception that the main and medical history questionnaires have been replaced by the new national questionnaire which covers both areas.

1. Paper Option

A follow-up package will be sent to those indicating interest containing an information and consent form outlining risks and benefits of the project. The information form will describe the questionnaires, physical measurements, donation of blood and urine samples, and will seek permission for the project to access their routinely collected health data and re-contact them in future. In addition it will request permission for use of all these data for cancer and chronic disease research by scientists.

The package will contain a bar coded paper copy of the baseline questionnaire with a request to complete and return it along with the signed consent form in a franked addressed envelope to the project. Even when the enrolment questionnaire is available for participants to complete online with full encryption protection, paper-based questionnaires will always be available for those who are uncomfortable with a computer. Since we may not be able to have Generations staff conduct physical measurements on new participants immediately, we will also provide a 72 inch

measuring tape for subjects to self-report waist and hip circumference, on the questionnaire, along with their weight and height.

Participants may also receive a laboratory requisition with a request to take this to a LifeLabs laboratory near their home or workplace to have blood drawn and a urine specimen taken. Blood and urine specimens will be taken by trained phlebotomists at LifeLabs Medical Labs Inc. into barcode numbered tubes and stored under refrigeration until the end of the day. All LifeLabs staff taking blood for the Project will be bound by the same confidentiality and privacy regulations as BC Generations Project staff. Each day the specimens will be securely transported to the BC Generations laboratory at the BC Cancer Research Centre where they will be processed and aliquoted into Nunc vials resulting in plasma, serum and DNA which can then be stored in -80°C freezers or liquid nitrogen. The bar coded tubes will only be linked to the participant through a reader that scans the barcode number into the participant's electronic file. In addition, to the separated blood products noted above, one tube of blood (ACD- yellow top) will be frozen at -80°C degrees to allow for immortalization of lymphocytes if needed to ensure a future supply of constitutional DNA from study participants. Specimens will be bar-coded with the participant's unique study number only. No personal identifying information will be present on the tubes. The specimen collection data, with participant unique study ID, will be stored on a separate server from the questionnaire data. The same electronic key used with questionnaire data will enable linkage with the participant's identity. The key as noted will be stored under password protection in an encrypted fashion on a third server.

2. On-line Option

2.1 Getting started:

After receiving an invitation letter, or learning about the BC Generations Project through a friend or through advertising, potential participants will contact us in the same ways as before, by phone or by email, to express their interest.

2.2 Identification:

The study clerk will email the potential participant a link to click in order to confirm his/her email address. This link is uniquely generated for each potential participant and takes him/her to a 'personal home page', where consent and questionnaire may be accessed. They will be asked to provide their name, email address, age, sex, postal code, a password to access the questionnaire website, the answer to a security question, and the answer to a communications question (that is, entering a distorted image to prevent automated responses). Once the potential participant's email address is confirmed, the study will email his/her enrolment ID, a copy of the consent form, a link to the study's main website, and the study's telephone number.

2.3 Consent:

The potential participant will need his/her enrolment ID and the password created during the registration phase to log in to the consent and questionnaire. The questionnaire link is not active until the consent process is completed. Participants will have the opportunity to review the consent form prior to logging on to the questionnaire website. In addition to being available on the BCGP main website, the consent form will be included in the email that contains the participant's login information. Additional information about the study is available on the website, and participants have the option of calling the BCGP call centre if additional information or clarification is required. For the consent process, the potential participant clicks on the link, and is presented with the full text of the BC information brochure and informed consent document (current approved version), and must click to advance through the pages, and then to accept it. After informed consent is provided, participants can then go on to complete the questionnaire at their leisure.

2.4 Questionnaire:

Participants will have the option of completing the questionnaire in English or French. Participants do not have to complete the questionnaire all at once; they may pause and come back to it at a later time, with 30 days to complete it. On completion, a thank-you email will be sent. Reminder emails will be sent 7, 14, and 28 days after participants login to start the questionnaire. After 30 days, their account will become inactive, and they will receive an email thanking them for their interest, and with instructions on how to re-activate their account if they choose.

Current BC Generations Project Assessment Centre Protocol

A small clinical assessment centre is running at the BC Cancer Research Centre. In addition to this assessment centre in Vancouver, the Project hopes to schedule assessment centres in major cities throughout British Columbia.

At entry participants review and sign the paper Participant Consent Form which includes permission to access past and future health information, store specimens and information and re-contact them in the future to request further information. At the assessment centre, participants will undergo a set of non-intrusive baseline measurements which will include some or all of height, weight, body fat, respiratory function, blood pressure, grip strength, and bone density. Participants may also be requested to complete the baseline information questionnaire, and donate blood and urine, or they may be asked to do these at a later time.

The Baseline Questionnaire—What Information do we Want?

The CPT questionnaire can be categorized into the following broad areas of interest:

Personal history of cancer and other diseases

Medical history, and general health questions, such as self-reported disabilities, as well as some limited phenotype information (related to skin and hair colour, chronic pain and chest pain, wheeze), will be collected using standardized questions adapted from those used successfully in various health surveys. These factors are important in any analysis examining health outcomes, because they may serve as predictors of future disease. Associations of *in utero* and early childhood exposures with common diseases of adult life have been widely reported.

Female reproductive health

Information on age at menstruation, age at first birth, number of children, hysterectomy, etc. have been shown to affect risk of a number of cancers, and presence of this data will enable control for known risk factors in investigating the effect of new putative causes of disease.

Family history of disease

Family history is a known predictor of common cancers, cardiovascular diseases, and a number of other medical conditions. Consequently, questions are included about the history of close family members (first degree relatives). As well being a member of twins or other multiple order birth could potentially identify subgroups of interest for future family-based studies.

Life habits and behavior

Lifestyle factors such as cigarette smoking and alcohol consumption are widely known to be associated with chronic diseases. Degree of physical activity, dietary patterns, etc. are also known or suspected to be related to various disease states. In designing questions, attention is being given to those questions which are likely to be reported reliably, are simple to answer, and give a relatively wide range of responses. In addition the questions are non-contentious in order to obtain as complete information as possible. The questions are important not only as potential risk factors in themselves but also as variables that must be controlled for in investigating new hypotheses.

Physical environment including residential history and usual occupation

These will include current address, residence at birth, and a residence history. Current address will allow researchers to explore multiple potential environmental risk factors by linking records anonymously with various Canadian national ecological databases (while maintaining confidentiality). For instance, residence information will allow determination of whether cancer is more common in areas close to known toxic waste sites. Occupational data (longest job; jobs held) will be collected and coded using Canadian Standard Occupation Coding manuals and Standard Industrial Classification manuals. This will allow the ability to explore the relevance of occupation as a socioeconomic and environmental determinant of disease. In addition, the blood and urine samples will allow a degree of quantification of a number of environmental and occupational exposures (such as organic chlorine compounds, heavy metals) which might, in conjunction with data on the questionnaire, help elucidate the etiologies of a number of chronic neurologic diseases as well as cancers.

Socio-demographics

Socioeconomic position and demographic markers are known to be correlated with mortality, measures of morbidity, and access to health services (White et al, 2007; Leon and Wilkinson 1989; Saposnik et al, 2008) Hence assessment of these factors, both as potential exposures and as confounders, is necessary for any longitudinal study. Health behaviors (such as physician, dentist use, screening program use; etc.) are very important in determining 'health consciousness' of subjects and speak to social factors (locus of control; connectedness) involved in predicting morbidity and mortality from disease.

Diet

Epidemiologic investigations and randomized trials have provided conflicting evidence regarding the effects of various dietary components (such as fat and fibre) on important disease outcomes (Bingham et al, 2003; Beresford et al, 2006; Howard et al 2006) and about the most appropriate method to approach measurement (Day et al, 2001; Willett, 2001; Schatzkin et al, 2003). The availability of biological samples in the CPFT cohort will allow the direct measurement of the levels of many biomarkers of interest (e.g. lipid profile, vitamins, red cell fatty acids). However, since biomarkers may not necessarily reflect actual intakes (Cade et al, 2002) and are not available for many dietary items, questionnaire methods must also be employed. Our approach is likely to be somewhat different from that used previously in studies in that we will seek a 3 day food diary from selected participants, which they can complete and return to us at their own convenience. The Alberta component of the Canadian Cancer Cohort will conduct a pilot project using the 3 day diary to assess its feasibility prior to the BC cohort making a decision on this factor.

Physical activity

The questions on physical activity that have been included in the cohort questionnaire were adapted, based upon piloting, from validated survey instruments used in previous community health surveys. They are principally intended to allow participants to be ranked according to their levels of physical activity (vigorous, moderate, and walking).

New environmental factors

A separate task force within the Canadian cancer cohort, headed by Dr. Jack Siemiatycki (Université de Montréal) has been constituted to examine and report back on strategies to assess how environmental exposures to potentially carcinogenic chemicals and other physical agents affect risk of cancer, heart, neurological, and other diseases. This committee will suggest cutting-edge methods for evaluating such exposures at the end of 2009. This will mean that participants may need to be re-contacted and re-consented for new tests.

Physical Measurements –What data do we want to collect?

Physical measures can provide important information on risk of development of chronic diseases. Bringing the initial 4,000 cohort participants into the assessment centre allowed the opportunity for measures to be made uniformly by professional staff, contributing to a high-quality research resource. The BC Generations Project measured weight, standing and sitting height, waist and hip circumference, blood pressure,

bioelectrical impedance, grip strength, spirometry, and bone densitometry. All staff conducting physical measurements were trained health technicians. Under the new protocol we will continue to ask for permission to have BC Generations Project staff collect measurement data, however, participants not attending an assessment centre will be notified that this will not occur immediately, and that they will be re-contacted in the future for this step. Participants will, of course, be free to refuse this step at re-contact. Participants attending assessment centres will have some or all of the physical measures described below measured.

Weight:

Weight is critical in determining body mass index, a commonly used measure of obesity. Overweight and Obesity (BMI >25) have been associated with a number of adverse health outcomes including diabetes, heart disease and cancer (Rexrode et al, 1997; Carey et al, 1997; Felson et al, 1988; Calle et al, 2003; IARC, 2002; den Tonkelaar et al, 1995; Giovannucci et al, 1995). Weight was measured at our original assessment on a Tanita BC-418 bioelectrical impedance instrument at the same time that body composition is measured. Under the new protocol, participants will be asked to self report weight. In the event that we are able to find funds for the mobile measurement unit weight will be recorded again by trained Generations staff.

Height:

Height measurement is also a critical variable in calculating body mass index, and was measured in initial participants using a Seca 214 Portable Stadiometer. Participants at the assessment centre were required to remove their shoes prior to measurement. Sitting height was also be measured with the participant seated on a standard chair. In the new protocol we will ask for self-reported standing height, but will have standing and sitting height re-measured by staff if possible

Waist and hip circumference:

Waist/hip ratio has been shown to be predictor of risk myocardial infarction (Yussuf et al, 2005) and for breast cancer mortality in women (Borugian et al, 2003) and may also be associated with prostate cancer risk in men. Accumulation of abdominal body fat may predict risk of a number of adverse health outcomes in the absence of obesity. Measures were initially taken using a Seca 200 measuring tape at the assessment centre. Participants were required to remove bulky clothing (for instance winter coats) to facilitate measurement. We will provide a measuring tape in the enrollment package for participants to self-report these parameters as we recruit the remaining 36,000 participants, but will have staff re-measure these in future if possible.

Bioelectrical impedance:

Bioelectrical impedance will provide an indication of proportional body composition (lean body mass vs fat). The measurement will be carried out by having participants remove shoes and socks, stand on the foot pads of a Tanita BC-418 instrument, and grasp both handles of the machine. Resistance to a small electrical current generated by the machine will be recorded as it travels through the participant's body. The amount of

resistance is inversely proportional to the amount of fat-free or lean body mass, as lean mass conducts electricity faster than fat. Fat mass (to the nearest 0.2 kg), fat free mass (to the nearest 0.2 kg), and total body water will be recorded for each arm, each leg, and the body trunk, giving information on fat distribution. In addition, impedance (to the nearest ohm) will be recorded. The measurement technician will ask whether the participant is pregnant or has a pacemaker and, if so, this measurement will not be done.

Bone densitometry:

A single measure of calcaneal bone density will be undertaken on the left heel using a GE Lunar Achilles Ultrasound Unit with the participant sitting upright. The measurement takes 1-2 minutes and requires that the shoe and sock be removed from the left foot. If a participant has a sore or wound on the heel of the left foot, the right foot will be assessed. Bone density is a predictor of risk of fall and of fracture, a major cause of morbidity and mortality in older people.

Blood pressure:

Blood pressure is known to predict risk of heart disease and peripheral circulation as well as dementia (Prospective studies collaboration, 2002; Qiu et al, 2005). Systolic and diastolic blood pressure will be recorded using the Omron HEM-907XL blood pressure monitor. Participants will be required to roll up the sleeves of long sleeve shirts or blouses for this measure. The measure will be taken twice over a period of 3 minutes with the participant comfortably seated.

Grip strength:

Hand grip strength is a predictor of all-cause and cardiovascular mortality, as well as disability (Rantanen et al, 1999, 2000; 2003; Metter et al, 2002). In addition, a study of European men and women found that low grip strength was associated with lower bone mass and, in women, with an increased risk of fracture (Dixon et al, 2005). Grip strength will be measured and recorded in kg for the right and left hand using the Digital Hydraulic Hand Dynamometer.

Spirometry:

Spirometry provides a good indicator of lung function, and also has been related to all cause mortality, cardiovascular and cerebrovascular disease, and self-reported health (Ebi-Kryston, 1988; Truelsen et al, 2001; Sin et al, 2005; Strachan 1992; Canoy et al, 2004; Myint et al, 2005). Forced expiratory volume over 1 second and total expiratory volume will be measured using a MiniSpir spirometer. Two quantitatively similar measurements are needed and will be taken over a time period of 6 minutes in order to allow participants recovery time. A maximum of 3 'blows' will be requested in order to obtain the 2 measures. A BC Generations Project staff member at the assessment centre checked that the participant did not have any contraindications to spirometry, such as a recent chest infection, previous heart attack, recent chest surgery, pneumothorax, abdominal or eye surgery, or history of detached retina. Spirometry will not be

undertaken in these participants. These same safeguards will be in place as we obtain spirometry on as many as possible of the remaining 36,000 participants.

Blood Specimen

The phlebotomist will check whether the participant has had any previous problems giving blood and will then inspect the suitability of the veins in the inner elbow region. If these veins appear suitable for blood collection, then this will be undertaken from the inner elbow (left or right as the participant chooses) using an 18G needle and barrel. Vacutainer tubes, which will be barcoded, will be used to collect blood. Blood will be collected into several types of tubes in a specific sequence to reduce errors and also to prioritize specimens in the event that a participant's vein is unable to sustain collection of the entire 40 ml (for instance, a participant undergoing chemotherapy).

Blood Processing

Processing of blood and urine samples at LifeLab facilities under the revised protocol will be minimal. As blood is collected from a participant, the vacutainers will be inverted ten times to mix the anticoagulant/preservative/clot activator with the whole blood. In addition, the blood in the serum separation tube will be centrifuged at 2500g for 10 minutes in a non-refrigerated centrifuge after clotting for 25-30 minutes. Specimens will be refrigerated and shipped to the BC Generation blood processing lab daily.

Once shipped to the BC Generations laboratory in the BC Cancer Agency, the unique barcode on each one will be scanned into the IT system. This is important to link the participant questionnaire data with the Laboratory Information Management System (LIMS). LIMS keeps track of what specimens have been provided by a participant, how they have been processed, and where they have been stored.

. Serum from each of 2 serum separator vacutainers will be divided into aliquots for storage. Each of 3 x 6 ml EDTA tubes will be centrifuged and pipetted into aliquots of plasma, buffy coat (for DNA extraction), and red cells.. The acid citrate dextrose tube will have DMSO added and will be stored in liquid nitrogen for future use as a source of lymphocytes. These will immediately be frozen down to be stored at -80°C or -160°C (liquid nitrogen) at our secure storage facility.

Urine Specimen

Each participant will be asked to provide a urine specimen at LifeLabs in a sterile plastic container, for the study of metabolites excreted from the body. The emerging field of metabolomics has demonstrated that proteins, peptides and other compounds in urine can help identify not only kidney and urinary tract diseases but also may serve as early markers for cancers and other diseases (Pisitkun et al, 2006; Barratt and Topham, 2007).

Although this science is still in its infancy, it shows great promise for advancing the discovery of early diagnostics for future use in medicine.

Participants will provide the specimen in a locked LifeLab washroom, place the plastic collection vessel in an opaque plastic bag, and deposit the bag in a collection rack at the door before leaving. Urine from the urine collection vessel will be transferred to a bar coded vacutainer by removing the protective label from the lid of the collection vessel and pushing the cap of the vacutainer onto the sheathed needle in the vessel recess.

The urine specimen will be pipetted into aliquots at the BC Generations lab, frozen down and transferred to secure storage at -80°C .

Saliva Specimen

In the future, we may wish to request a saliva sample, for its own sake or in those instances where a participant cannot give a blood sample due to contraindications or location. In these instances, a saliva sample will be obtained using the Oragene kit and stored according to kit instructions.

Long Term Storage of Biospecimens

Blood and urine specimens will be tracked in storage in 2 ways. As noted earlier, all aliquots will be bar coded and, secondly, they will be stored in a -80°C freezer or in vapour phase liquid nitrogen (LN) (-160°C) according to a positional numbered grid maintained on the Laboratory Information Management System (LIMS). Thus all specimens will be able to be positively linked to the participant's unique study number and hence to his/her information without use of any personal identifying information.

The freezers and vapour phase liquid nitrogen containers will be stored at the BC Cancer Research Centre or in a secure locked facility at the BC Genome Sciences Centre (a part of the BC Cancer Agency) . These facilities have on-site security during working hours, and access to floors above the main at the BC Cancer Research Centre is controlled by individually coded proximity cards. Finally, the specimen storage room within the laboratory will be kept locked.

Each freezer and vapor phase LN storage tank will be connected to an alarm system, which monitors temperature and alerts security to possible failure. Each freezer is designated to a lab technician who is responsible for responding to any alerts generated by the alarm system. A further technician is designated as the secondary responder, with the principal investigator as the tertiary responder. Names and work as well as home phone numbers are clearly indicated on the front door of each freezer/LN tank. In the event of an alarm, security personnel move through the designated response personnel until contact is made. In the case of an evening or night alert, technicians respond by coming to the research centre, diagnosing the problem, and taking appropriate action. In the case of the LN storage tanks, the specimens are stable to a temperature of -160°C for

at least 7 days allowing plenty of time for repairs to be effected. In the event of a compressor failure in a -80°C freezer (the most serious problem), the technician will have about 3 hours to transfer specimens to a new freezer. The BC Generations Project has arranged to have 1 empty freezer running at all times as a back-up in case of a compressor failure. Agreements with scientific equipment vendors provide for the interim replacement freezer to be loaned to the cohort while repairs are effected on a defective machine. Thus there should never be more than 24 hours in which an empty back-up freezer is not available in immediate proximity to the stored specimens.

Short-Term Storage of Participant Information

During conduct of the recruitment phase of the cohort, data will be stored on servers at the BC Cancer Agency's Cancer Control Research Program. The mailing and appointment database will be maintained on a password protected server, and access will be granted only to named appointment clerks and programmers. Each log-on to the database will result in an audit trail delineating time of log-on and log-off and the name of the staff member seeking access.

Questionnaire responses and measurement results will be collected using their unique study identification number allocated at invitation. This information will be maintained on a second server, with independent password and user access trails. Further, the key which allows linkages between the name, address, and phone number of each participant and his/her unique cohort identification number will be resident on the third server, again with independent password and encryption protection, and audit trail capability.

In the new on-line process, when participants register their interest in the study, two separate identification numbers are created: the "enrolment ID" is attached to the participant's personally identifying information (e.g., name, contact information, medical card number). The "participant ID" is attached to all de-identified data (e.g., most questionnaire responses). This will ensure that the identifier attached to most of the study data is distanced from the identifier attached to personally identifying data. The enrolment ID and the participant ID will not contain any identifying information (e.g. date of birth). Personally identifying information will be stored in a separate database than de-identified data. All databases will be encrypted and data transfer will only occur over secure networks.

Information collected from participants using web-based tools will be transferred over a secure network to highly secure, multiply redundant and encrypted firewall-protected servers. All data will be saved in encrypted, password-protected files. Only authorized staff will have access to the study data. The servers will be housed in a physically secure environment requiring key card access.

Long Term Storage of Participant Information

Because of the sensitivity of the information collected from participants, the BC Generations Project is negotiating to store participant data, participant identifying information, and participant key file linking ID and study number long-term (beyond the recruitment phase) at the Population Data BC facility at the University of BC. This facility housed at the University of BC has established strict parameters for data security, including a locked physical plant within which all the facility's computers are kept, further locked doors within the plant (Red zone) within which all personal identifying information is kept, and access to the 'red zone' restricted to a small number of named computer programmers. Sign-on to any computers in Population Data BC leaves an audit trail linked to an individual programmer or investigator. Furthermore, computers within the personal information or 'red zone' use a different set of servers than those on which health data are stored. The system does not permit any data carrying personal identifying information to be released outside the 'red zone.' Anonymized data abstracts released to bona fide investigators are held in a distinct workspace cut off electronically from the red zone and from the servers carrying health data. Investigators conduct analyses using virtual personal networks. Additional details are available in the privacy impact statement on the Population Data BC website <http://www.popdata.bc.ca/privacy>

This system, recently reviewed by the Office of the BC Information and Privacy Commissioner, is operational. Discussions are currently underway with Population Data BC (Nancy Meagher, Executive Director) on the nature of the relationship to be established with the BC Generations Project. The most likely relationship will be to frame Population Data BC as a service provider to the Generations Project; however, other closer relationships are also being discussed. The BC Cancer Agency will be the public body contracting for the services of Population Data BC.

Further Contact with Cohort Members

In order to maximize the value of the BC Generations Project over its useful lifetime, further contact will be requested with subsets of participants within the Cohort. Prior to any attempt to re-contact participants with a request for new tests or new information, an amendment will be submitted to the Research Ethics Board. In order to facilitate this follow-up activity, a separate copy of the identifiers, information, and linkage key will be made with each submission of data to Population Data BC on an encrypted flash drive, which will be stored in a secure vault by Iron Mountain Ltd. This drive will be password protected and will be used only by cohort staff within the BC Cancer Agency to re-contact participants. Collection of new information, or updating of older information, will proceed in the same fashion as outlined above in the section on short-term storage of participant information and, at the conclusion of data acquisition, the encrypted files will once more be returned to Population Data BC, with a copy on a flash drive stored in a secure vault.

All re-contact with participants will be in the form of a written communication asking for either further information on aspects of their lifestyle (for instance, diet) or their

environmental exposures (for instance to sunlight, occupation). Actual data collection may be done using computer-aided telephone interview (CATI), online questionnaire or postal instruments by BC Generations Project staff, who will be BC Cancer Agency employees. Re-contact might also request donation of a further specimen (for instance, fingernail clippings or a lock of hair for assessing body burden of heavy metals) as scientific knowledge advances.

It is our intention to continue to communicate with BC Generations Project participants on an annual basis in any case. This will likely take the form of a newsletter delivered in hard copy or, if preferred, in electronic form. The intention of this communication will be to provide the latest scientific information emerging from the CPFT and also to provide topical health advice. In order to facilitate maintenance of communication with BC Generations Project members, participants will be asked for the name, address, and phone number of a close personal friend who will always know the whereabouts of the member should a move take place. This is a standard practice in longitudinal surveys.

References

Barratt J, Topham P (2007) Urine proteomics: the present and future of measuring urinary protein components in disease. *CMAJ* 177:361-368.

Beresford SAA, Johnson KC, Ritenbaugh C, et al. Low-fat dietary pattern and risk of colorectal cancer: the Women's Health Initiative Randomized Controlled Dietary Modification Trial. *JAMA* 2006; 295: 643-54.

Bingham S, Luben R, Welch A, Wareham NJ, et al. Are imprecise methods obscuring a relation between fat and breast cancer? *Lancet* 2003; 362: 212-14.

Borugian MJ, Sheps SB, Kim-Sing C, Olivotto IA, van Patten C, Dunn BP, Coldman AJ, Potter JD, Gallagher RP, Hislop TG (2003) . Waist-to-hip ratio and breast cancer mortality. *Am. J. Epidemiol*;158:963-968.

Cade J, Thompson R, Burley V, et al. Development, validation and utilization of food frequency questionnaires - a review. *Public Health Nutr* 2002; 5: 567-87.

Calle EE, Rodriguez C, Walker-Thurmond K et al (2003). Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *N Engl J Med*; 348:1625-38.

Canoy D, Luben R, Welsh A, et al (2004). Abdominal obesity and respiratory function in men and women in the EPIC-Norfolk study, United Kingdom. *Am J Epi*; 159:1140-49.

Carey VJ, Walters EE, Colditz GA et al (1997). Body fat distribution and risk of non-insulin dependent diabetes mellitus. *Am J Epi* ; 145:614-19.

Cobiak (2007) Epigenomics and nutrition. *Forum Nutr.* 60:31-41.

Collins FS (2004) The case for a US prospective cohort study of genes and environment. *Nature* 429:475-477.

Courneya KS, Friedenreich CM. Physical activity and cancer control *Semin Oncol Nurs* 2007; 23:242-252.

Day NE, McKeown N, Wong MY, et al. Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium. *Int J Epi* 2001; 30:309-17.

den Tonkelaar I, Seidell JC, Collette HJA (1995). Body fat distribution in relation to breast cancer in women participating in the DOM-project. *Breast Cancer Research and Treatment*; 34:55-61.

Dixon WG, Lunt M, Pye SR, et al (2005). Low grip strength is associated with bone mineral density and vertebral fracture in women. *Rheumatology*; 44:642-46.

Ebi-Kryston KL (1988). Respiratory symptoms and pulmonary function as predictors of 10-year mortality from respiratory disease, cardiovascular disease and all causes in the Whitehall Study. *J Clin Epi* ; 41:251-60.

Felson DT, Anderson JJ, Naimark A et al (1988). Obesity and knee osteoarthritis: The Framingham Study. *Ann Intern Med*; 109:18-24.

Freedman LS, Potishman N, Kipnis V, et al. A comparison of two dietary instruments for evaluating the fat-breast cancer relationship *Int J Epidemiol* 2006 35:1011-1021.

Giovannucci E, Ascherio A, Rimm EB, et al (1995). Physical activity, obesity and risk for colon cancer and adenoma in men. *Ann Intern Med*; 122:327-34.

Harvard Center for Cancer Prevention (1996) Harvard report on cancer prevention Vol 1, causes of human cancer. *Cancer Causes Cont.* 7: Suppl 1:S3-59.

Hoover RN (2007) The evolution of epidemiologic research from cottage industry to big science *Epidemiology* 18:13-17.

Howard BV, Van Horn L, Hsia J, et al. Low-fat dietary pattern and risk of cardiovascular disease: the Women's Health Initiative Randomized Controlled Dietary Modification Trial. *JAMA* 2006; 295:655-66.

Hunter DJ, Thomas G, Hoover RN, Channock SJ (2007) Scanning the horizon: what is the future of genome-wide association studies in accelerating discoveries in cancer etiology and prevention? *Cancer Causes Control* 18:479-484.

IARC (1992) IARC monographs on the evaluation of carcinogenic risks to humans vol 55, solar and ultraviolet radiation. International Agency for Research on Cancer, Lyon, pp 1-316.

IARC (2001) IARC handbooks of cancer prevention vol 5, sunscreens. Lyon, pp1-93.

IARC Handbooks of Cancer Prevention (2002) : Volume 6, Weight Control and Physical Activity IARC, Lyon.

International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921.

Jones PA, Baylin SB (2007) The epigenomics of cancer. *Cell* 128:683-692.

Kristal AR, Peters U, Potter JD Is it time to abandon the food frequency questionnaire *Cancer Epidemiol Biomark Prev.* 2005; 14:2826-2829.

Le Fanu (1999) *The Rise and Fall of Modern Medicine.* Little-Brown, New York.

Mathers JC (2007) Early nutrition: Impact of Epigenetics. *Forum Nutr.* 60:42-48.

Metter EJ, Talbot LA, Schrage M, et al (2002). Skeletal muscle strength as a predictor of all-cause mortality in healthy men. *J Gerontology Series A Biological Sciences and Medical Sciences*; 57A:B359-65.

Morton LM, Turner JR, Cerhan JR *et al* (2007) Proposed classification of lymphoid neoplasms for epidemiologic research from the pathology working group of the international lymphoma epidemiology consortium (InterLymph) *Blood* 110:695-708.

Myint PK, Luben RN, Surtees PG, et al (2005). Respiratory function and self-reported functional health: EPIC-Norfolk population study. *Eur Resp J*; 26: 494-502.

Neilson HK, Robason PJ, Friedenreich CM, Csizmadi I . Estimating activity energy expenditures: how valid are physical activity questionnaires. *Am J Clin Nutr* 2008; 87:279-291.

Pisitkun T, Johnstone R, Knepper MA (2006) Discovery of urinary biomarkers. *Mol. Cell Proteomics*; 5:1760-1771.

Potter JD (2004) Toward the last cohort. *Cancer Epidemiol Biomarkers Prev.* 13:895-897.

Prospective Studies Collaboration (2002) Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 360:1903-1913.

Qui C, Winblad B, Fratiglioni L, (2005) The age-dependent relation of blood pressure to cognitive function and dementia . *Lancet Neurol* 4:487-499.

Rantanen T, Guralnik JM, Foley D, et al (1999). Mid life hand grip strength as a predictor of old age disability. *JAMA*; 281:558-60.

Rantanen T, Harris T, Leveille SG, et al (2000). Muscle strength and body mass index as long term predictors of mortality in initially healthy men. *J Gerontology Series A Biological Sciences and Medical Sciences*; 55A:M168-73.

Rantanen T, Volpato S, Ferrucci L, et al (2003). Hand grip strength and cause-specific and total mortality in older disabled women: exploring the mechanism. *J Am Geriatrics Society*; 51:636-41.

Rees JL (2004) The genetics of sun sensitivity in humans. *Am J Hum Genet* 73:739-751.

Rexrode KM, Hennekens CH, Willett WC et al (1997). A prospective study of body mass index, weight change, and risk of stroke in women. *JAMA*; 277:1539-45.

Rhodes R (1986). *The Making of the Atomic Bomb*. Simon and Schuster, New York.

Rothman N, Skibola CF, Wang SS, *et al* (2006) Genetic variation in TNF and IL-10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium. *Lancet Oncol* 7:27-38

Saposnik G, Cote R, Phillips S et al (2008) Stroke outcome in those over 80: a multicentre cohort study across Canada *Stroke* 39:2310-2317.

Schatzkin A, Kipnis V, Carroll RJ, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. *Int J Epi* 2003; 32:1054-62.

Schouten LJ, Rivera C, Hunter DJ, et al.(2008) Height body mass index and ovarian cancer: a pooled analysis of 12 cohort studies. *Cancer Epidemiol, Biomark and Prev* 17:902-912.

Sin DD, Wu LL, Man SFP (2005). The relationship between reduced lung function and cardiovascular mortality: a population-based study and a systematic review of the literature. *Chest*; 127:1952-59.

Smith GD, Ibrahim S. (2001) Epidemiology-is it time to call it a day? *Int J Cancer* 30:1-11.

Strachan DP (1992). Ventilatory function, height and mortality among lifelong non-smokers. *J Epi Comm Health*; 46:66-70.

Suzuki T, Matsuo K, Hasegawa Y et al. Anthropomorphic factors at age 20 years, and risk of thyroid cancer. *CCC* 2008; Jul 10 Epub ahead of print.

Terry PD, Goodman M (2006). Is the association between cigarette smoking and breast cancer modified by genotype? A review of epidemiologic studies and meta-analysis. *CEBP* 15:602-611.

Truelsen T, Prescott E, Lange P, et al (2001). Lung function and risk of fatal and non-fatal stroke. The Copenhagen City Heart Study. *Int J Epi*; 30:145-51.

Yusuf S, Hawken S, Ôunpuu S et al (2005). Obesity and the risk of myocardial infarction in 27000 participants from 52 countries: a case-control study. *Lancet*; 366:1640-9.

Venter JC *et al.* (2001) The sequence of the human genome. *Science* 291:1304-1351.

Wang SS, Slager SL, Brennan P *et al* (2007). Family history of hematopoietic malignancies and risk of non-Hodgkin lymphoma (NHL): a pooled analysis of 10,211 cases and 11,905 controls from the international lymphoma epidemiology consortium (InterLymph). *Blood* 109:3479-3488.

Weidman JR, Dolinoy DC, Murphy SK, Jirtle RL (2007) Cancer susceptibility: epigenetic manifestations of environmental exposures. *Cancer J.* 13:9-16.

White C, Glickman M, Johnson B, Cobin T (2007) Social inequalities in adult male mortality by the National Statistics socio-economic classification, England and Wales 2001-03 *Health Stat Q*; Winter (39):6-23.

Willett WC. Commentary: Dietary diaries versus food frequency questionnaires - a case of indigestible data. *Int J Epi* 2001; 30:317-19.