

Research article

Open Access

The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer

Esther M John¹, John L Hopper², Jeanne C Beck³, Julia A Knight⁴, Susan L Neuhausen⁵, Ruby T Senie⁶, Argyrios Ziogas⁵, Irene L Andrulis⁷, Hoda Anton-Culver⁵, Norman Boyd⁸, Sandra S Buys⁹, Mary B Daly¹⁰, Frances P O'Malley¹¹, Regina M Santella⁶, Melissa C Southey¹², Vickie L Venne⁹, Deon J Venter¹², Dee W West¹, Alice S Whittemore¹³, Daniela Seminara¹⁴ for the Breast Cancer Family Registry

¹Northern California Cancer Center, Union City, California, USA

²Center for Genetic Epidemiology, The University of Melbourne, Victoria, Australia

³Coriell Institute for Medical Research, Camden, New Jersey, USA

⁴Division of Epidemiology and Biostatistics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada

⁵Division of Epidemiology, Department of Medicine, University of California Irvine, Irvine, California, USA

⁶Mailman School of Public Health of Columbia University, New York, New York, USA

⁷Fred A. Litwin Center for Cancer Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada

⁸Division of Epidemiology and Statistics, Ontario Cancer Institute, University Health Network, Toronto, Ontario, Canada

⁹Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah, USA

¹⁰Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA

¹¹Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada

¹²Department of Pathology, The University of Melbourne, Victoria, Australia

¹³Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California, USA

¹⁴Clinical and Genetic Epidemiology Research Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland, USA

Corresponding author: Daniela Seminara, seminard@epndce.nci.nih.gov

Received: 13 Jan 2004 Revisions requested: 9 Feb 2004 Revisions received: 5 Apr 2004 Accepted: 19 Apr 2004 Published: 19 May 2004

Breast Cancer Res 2004, **6**:R375-R389 (DOI 10.1186/bcr801)

© 2004 John *et al.*; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Introduction The etiology of familial breast cancer is complex and involves genetic and environmental factors such as hormonal and lifestyle factors. Understanding familial aggregation is a key to understanding the causes of breast cancer and to facilitating the development of effective prevention and therapy. To address urgent research questions and to expedite the translation of research results to the clinical setting, the National Cancer Institute (USA) supported in 1995 the establishment of a novel research infrastructure, the Breast Cancer Family Registry, a collaboration of six academic and research institutions and their medical affiliates in the USA, Canada, and Australia.

Methods The sites have developed core family history and epidemiology questionnaires, data dictionaries, and common protocols for biospecimen collection and processing and pathology review. An Informatics Center has been established to collate, manage, and distribute core data.

Results As of September 2003, 9116 population-based and 2834 clinic-based families have been enrolled, including 2346 families from minority populations. Epidemiology questionnaire data are available for 6779 affected probands (with a personal history of breast cancer), 4116 unaffected probands, and 16,526 relatives with or without a personal history of breast or ovarian cancer. The biospecimen repository contains blood or mouthwash samples for 6316 affected probands, 2966 unaffected probands, and 10,763 relatives, and tumor tissue samples for 4293 individuals.

Conclusion This resource is available to internal and external researchers for collaborative, interdisciplinary, and translational studies of the genetic epidemiology of breast cancer. Detailed information can be found at the URL <http://www.cfr.epi.uci.edu/>.

Keywords: biospecimen repository, breast cancer, familial aggregation, genetic epidemiology

Introduction

Breast cancer is known to be 'familial', both from the clinical perspective of observing more families with multiple cases than would be expected by chance, and from the population perspective of an increased risk of breast cancer among women with a family history of the disease. The two-fold to threefold increased risk to first-degree relatives of affected women is likely to be due to underlying familial factors, both genetic and environmental (for example, hormonal and lifestyle factors), and the risk gradient across these factors must be 20–100-fold or more [1,2]. Understanding the interplay of genetic and environmental causes of familial aggregation is therefore a key to understanding the causes of breast cancer and to facilitating the development of effective prevention and therapy.

Much is yet to be learned about the causes of familial aggregation of breast cancer. Pathogenic mutations in the genes *BRCA1* [3] and *BRCA2* [4] are associated with large individual increased risks, on the order of 10–20-fold, but, being rare, they explain less than 20% of the increased risk associated with having an affected first-degree relative [1,5]. Less than half the families with three or more affected members in the Breast Cancer Linkage Consortium have segregating deleterious mutations in *BRCA1* or *BRCA2* [6]. Mutations in *TP53* [7], and possibly in the *ATM* [8] and *CHK2* [9] genes, seem to confer moderately increased risks of breast cancer but might explain only a very small proportion of familial aggregation. Breast cancer risk might also be affected by multiple variants in genes involved with hormonal or other etiological pathways, and the variants might be common and have small or modest effects on individual risk. Lifestyle and other known risk factors are unlikely to explain much familial aggregation [2]. The effects of some established lifestyle factors might vary in relation to family history [10] or *BRCA1* or *BRCA2* mutation status [11]. Thus, familial aspects of breast cancer are complex, potentially involving multiple genes, multiple environmental exposures, and 'gene–environment interactions' [12].

To address many unanswered research questions regarding the etiology of breast cancer and to expedite the translation of research results to affected and at-risk populations, the National Cancer Institute of the USA supported the establishment of a novel international research infrastructure for interdisciplinary and translational studies of the genetic epidemiology of breast cancer. The Breast Cancer Family Registry is a collaboration of six academic and research institutions and their medical affiliates located in the USA, Canada, and Australia. This paper describes the development of the Breast Cancer Family Registry research infrastructure, the resources available to the research community as of September 2003, and many of the possible studies using this resource.

Methods

Structure of the Breast Cancer Family Registry

The Breast Cancer Family Registry was established in 1995, with six participating sites from the USA, Canada, and Australia ascertaining families either from cancer registries (identifying population-based families) or seen in clinical and community settings (identifying clinic-based families) (Fig. 1). Population-based families were recruited from the Greater San Francisco Bay area, California, USA, by the Northern California Cancer Center; from the province of Ontario, Canada, by Cancer Care Ontario; and from the metropolitan areas of Melbourne and Sydney, Australia, by the University of Melbourne and the New South Wales Cancer Council. Clinic-based families, including those of Ashkenazi Jewish ancestry, were recruited from their local populations in the USA by Columbia University in New York City, New York, the Fox Chase Cancer Center in Philadelphia, Pennsylvania, and Huntsman Cancer Institute at the University of Utah in Salt Lake City, Utah; and in Australia by the University of Melbourne and New South Wales Cancer Council in Melbourne and Sydney, Australia. In Ontario, Canada, recruitment of clinic-based families was limited to Ashkenazi Jewish families.

The Breast Cancer Family Registry investigators include epidemiologists, molecular biologists, molecular geneticists, clinicians, geneticists, genetic counselors, statisticians, pathologists and behavioral scientists. The participating sites are supported through Cooperative Agreements; thus, the leadership and scientific conduct of the Breast Cancer Family Registry are a combined effort of the six principal investigators and their teams, with substantial involvement of the Program Officer and other representatives of the National Cancer Institute.

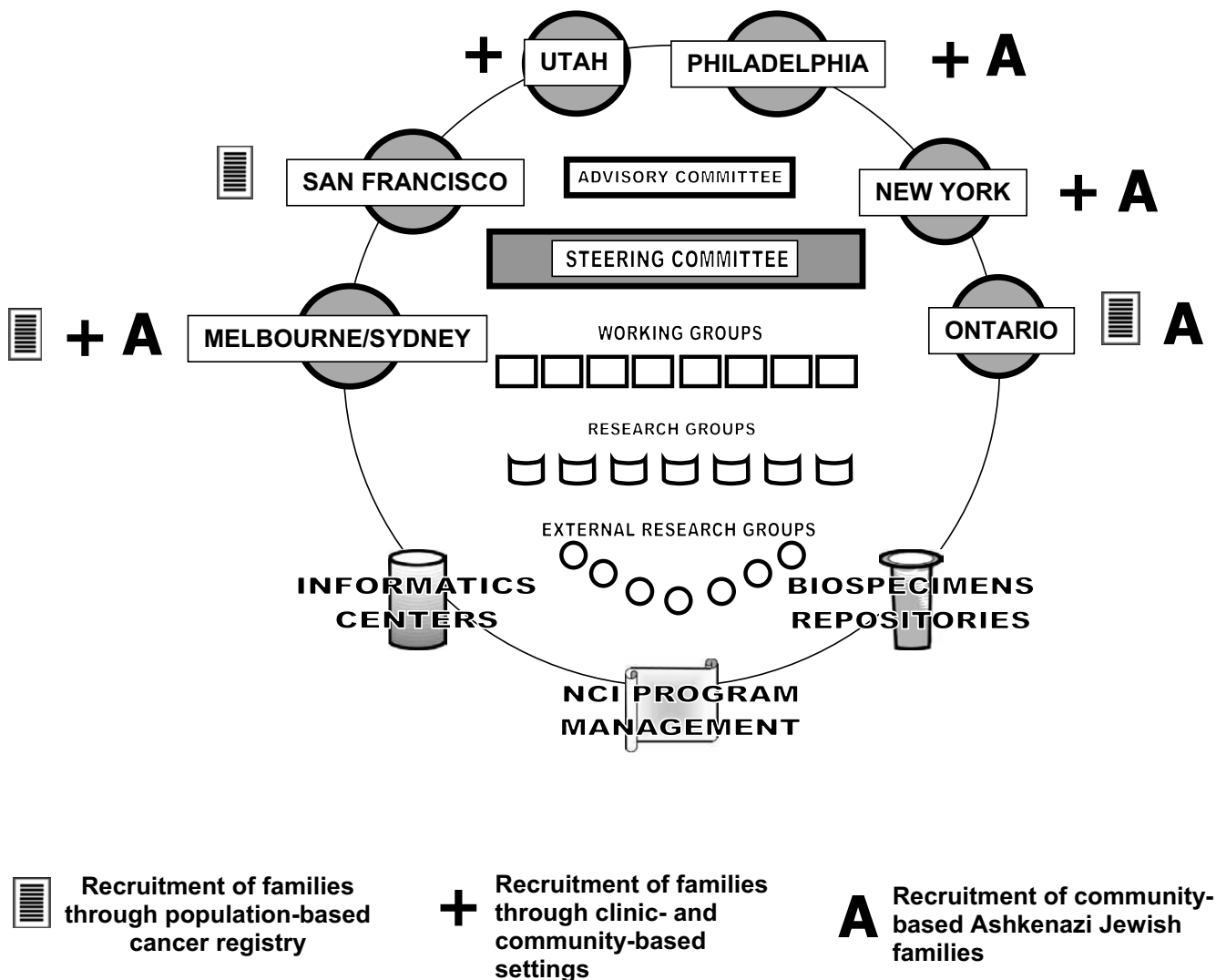
Policy and governance

The six sites have collaborated to develop and maintain the resources, to conduct interdisciplinary research, and to establish collaborations with external investigators. An organizational chart for the Breast Cancer Family Registry is provided in Fig. 2. Detailed information on the governance and policy can be found at the URL http://www.cfr.epi.uci.edu/nci/access_manual_05-29-02.htm. Research proposals from both internal and external investigators requesting access to the Breast Cancer Family Registry resources are evaluated by an Advisory Committee for scientific merit and the appropriate use of resources on the basis of criteria established by the Steering Committee of the Breast Cancer Family Registry. The decisions of the Advisory Committee are then reviewed and ratified by the Steering Committee.

Informatics Center

In 1998 an Informatics Center was established at the University of California, Irvine. The Informatics Center has

Figure 1



Structure of the Breast Cancer Family Registry.

developed a flexible and evolving informatics model, which maintains an Oracle relational database using a 'minimal data set' to track subjects. The Informatics Center receives standard data on several modules including family history, epidemiologic risk factors, diet, biospecimen tracking, genotyping, pathology, and follow-up data. The Informatics Center collates, manages, and distributes core data, in collaboration with each of the six local informatics units. Data from each site are submitted in batches to the Central Informatics System with the use of a secure access procedure. A quality assurance system ensures the reliability, validity, and completeness of the database. Up-to-date extract files are created from the relational database and are distributed to investigators. The Informatics Center also maintains the website for the Breast Cancer Family Registry <http://www.cfr.epi.uci.edu>, and coordinates teleconferences and other activities between the sites, the Advisory Committee,

the Steering Committee, the National Cancer Institute, and the Informatics Center.

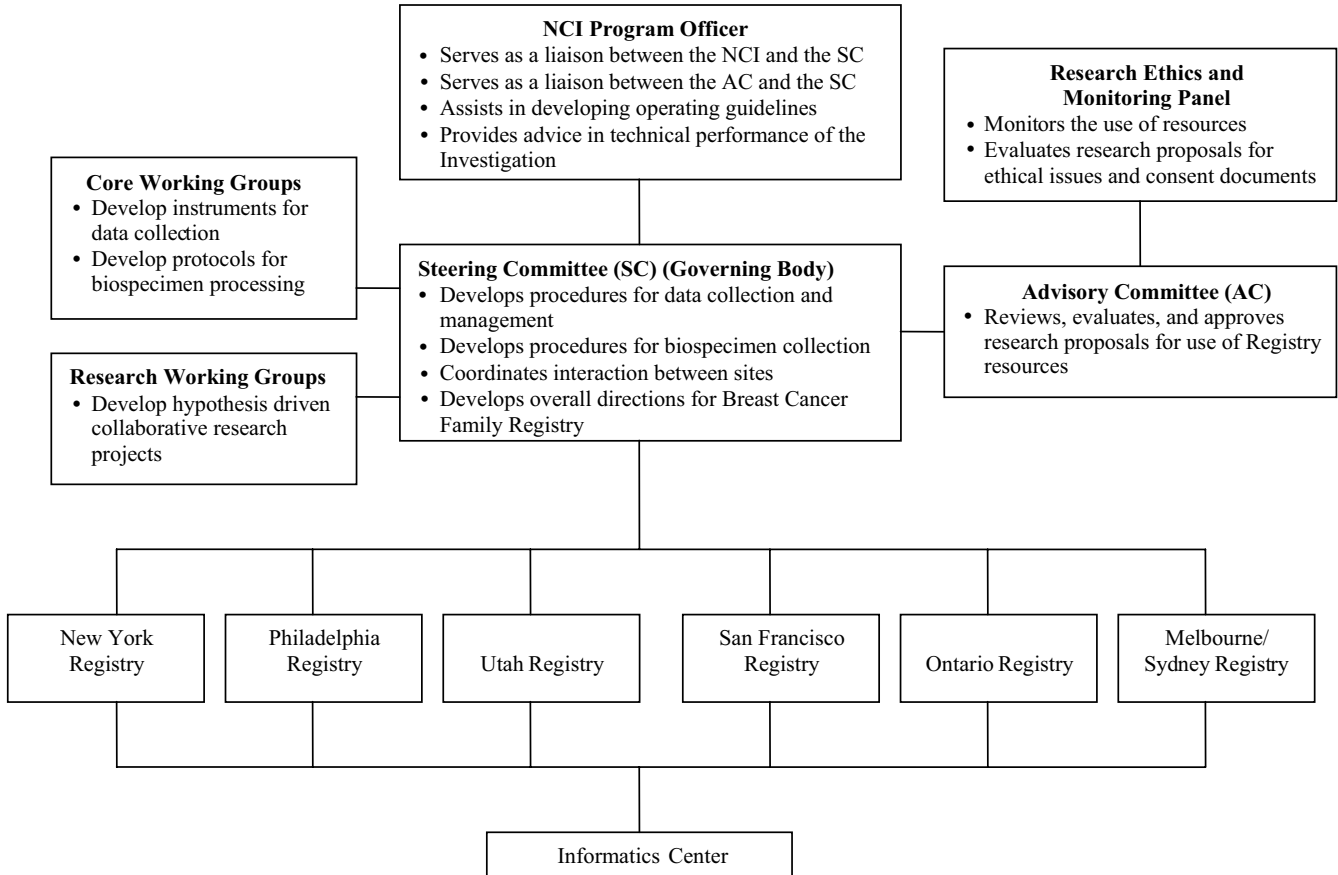
Ascertainment of probands and family members

Most families were enrolled in the Breast Cancer Family Registry from 1996 to 2000. During the period 2001–2005, several sites are continuing to recruit the following: (1) families known to segregate *BRCA1* or *BRCA2* mutations; (2) families with multiple cases of breast or ovarian cancer; (3) selected additional relatives of previously enrolled families; (4) families of Ashkenazi Jewish ancestry; and (5) families from specific racial and ethnic groups.

Clinic-based and community-based recruitment

Four sites enrolled families with multiple or early-onset cases of breast or ovarian cancer identified through community contacts and clinical settings including screening

Figure 2



Organization of the Breast Cancer Family Registry.

centers, family cancer clinics, surgical and medical oncology offices, and the Australian twin registry. Proband was defined as the first family member enrolled in the Breast Cancer Family Registry and may or may not have had a personal history of breast or ovarian cancer. Eligibility was based on one or more of the following criteria: two or more relatives with a personal history of breast or ovarian cancer; a woman diagnosed with breast or ovarian cancer at a young age; a woman with a history of both breast and ovarian cancer; an affected male; or known *BRCA1* or *BRCA2* mutation carriers. The Australian site also enrolled twin pairs in which one or both members had a personal history of breast cancer. Table 1 shows the eligibility criteria and family characteristics of clinic-based probands for each site.

Population-based recruitment

Three sites enrolled families through females with incident breast cancer identified through population-based cancer registries in defined geographic areas. Two sites also enrolled families through male breast cancer cases. Case probands, defined as the affected persons ascertained

from a cancer registry, were sampled according to one or more criteria, including age at diagnosis, gender, race/ethnicity, and family history. Table 2 provides an overview of the sampling strategies for the population-based case probands. Control probands were randomly sampled from the general population living in the relevant catchment area of each of the regional cancer registries, using random-digit dialing (San Francisco), lists of randomly selected residential telephone numbers (Ontario), and electoral rolls (Melbourne and Sydney). At all six sites, permission and assistance were sought from the proband to contact eligible relatives.

Special recruitment initiatives
Ashkenazi families

After the discovery that three specific mutations in *BRCA1* and *BRCA2* are relatively common among people of Ashkenazi Jewish ancestry [13], four sites (New York, Philadelphia, Ontario, and Melbourne and Sydney) were funded between 1996 and 2000 to recruit Ashkenazi Jewish families through their local communities and cancer family clinics, in addition to those being recruited through

Table 1

Ascertainment criteria for clinic-based families, including Ashkenazi Jewish, identified at five sites from 1996 to 2000

Parameter	Philadelphia ^a	New York ^a	Utah ^b	Melbourne and Sydney, Australia	Ontario, Canada ^c
Geographic area	Philadelphia	New York, New Jersey, Connecticut	Salt Lake City	Melbourne, Sydney	Province of Ontario
Recruitment criteria for families	Male with breast cancer; female with breast or ovarian cancer diagnosed at age <35; female with breast and ovarian cancer diagnosed at any age; female with bilateral breast cancer and first diagnosis at age <50; two individuals in two generations with breast or ovarian cancer; <i>BRCA1</i> or <i>BRCA2</i> mutation carrier	Male with breast cancer; female with breast or ovarian cancer diagnosed at age <45; female with breast and ovarian cancer diagnosed at any age; two first- or second-degree or three first-, second-, or third-degree relatives with breast or ovarian cancer; <i>BRCA1</i> or <i>BRCA2</i> mutation carrier	Male with breast cancer; female with breast or ovarian cancer diagnosed at age <45; female with breast and ovarian cancer diagnosed at any age; three individuals in two generations with breast or ovarian cancer, one diagnosed at age <45; <i>BRCA1</i> or <i>BRCA2</i> mutation carrier	Two or more first- or second-degree relatives with breast or ovarian cancer plus one or more of the following: additional relative(s) with breast or ovarian cancer, breast cancer diagnosed at age <40, ovarian cancer diagnosed at any age, bilateral breast cancer, both breast and ovarian cancer, breast cancer in a male relative	
Recruitment criteria for Ashkenazi Jewish families	One Jewish parent	Same criteria as above		Female with breast or ovarian cancer; one or more first- or second-degree relatives with breast or ovarian cancer	One Jewish grandparent
Relationship to proband	First- and second-degree relative	First-, second-, and third-degree relative	Any relative	Any relative	

^aRecruitment in 2001–2005 is continuing for minority families. ^bRecruitment in 2001–2005 is continuing for *BRCA1* and *BRCA2* mutation carrier families. ^cRecruitment in 2001–2005 is continuing for Ashkenazi Jewish families.

Table 2

Ascertainment criteria for probands of population-based families identified at three sites through local cancer registries from 1996 to 2000

Parameter	San Francisco, California ^a	Ontario, Canada ^b	Melbourne and Sydney, Australia
Geographic area	Greater San Francisco Bay Area	Province of Ontario	Metropolitan Areas of Melbourne and Sydney
Period of diagnosis	1995–1998	1996–1998	1996–1999
Sampling criteria for females	Age 18–34: include all Age 35–64: include all meeting high risk criteria ^c Age 35–64: sample 2.5% of whites and 15% of minorities not meeting high risk criteria	Age 18–35: include all Age 36–54: include all meeting high risk criteria ^d or Ashkenazi heritage; sample 25% of those not meeting high risk criteria Age 55–69: sample 35% of those meeting high risk criteria ^d or Ashkenazi heritage; sample 8.75% (35% × 25%) of those not meeting high risk criteria	Age 18–39: include all Age 40–49: 50% random sample Age 50–59: 25% random sample
Sampling criteria for males	Age 20–79: include all	Age 20–79: include all	Not applicable

^aRecruitment in 2001–2005 is continued for minority families from San Francisco Bay area (namely Latino, African-American, Chinese, Japanese, and Filipino families) and Orange County, California (namely Latino, African-American, and Filipino families). ^bRecruitment in 2001–2005 is continued for minority families. ^cBilateral breast cancer with first diagnosis before age 50 years, previous diagnosis of ovarian or childhood cancer, one or more first-degree relatives with breast, ovarian or childhood cancer. ^dPrevious diagnosis of breast or ovarian cancer; at least one first- or two second-degree relative(s) with breast or ovarian cancer; at least one second- or third-degree relative with breast cancer diagnosed before age 36 years, ovarian cancer diagnosed before age 61 years, multiple breast cancer primaries, both breast and ovarian cancer, or male breast cancer; at least three first-degree relatives with any combination of breast, ovarian, colon, prostate, or pancreatic cancer or sarcoma, with at least one diagnosis before age 51 years.

the 'core' recruitment activities described above. Individuals of Ashkenazi Jewish ancestry were also identified through the epidemiology or family history questionnaires.

Racial and ethnic minorities

To increase the racial and ethnic diversity of the resource, special efforts were undertaken at several sites in the USA to enroll African-American, Asian, and Hispanic families, either through oversampling of probands (for example, in San Francisco) or community outreach (for example, in New York). Recruitment of African-American, Asian, and Hispanic families is continuing, and in California it has been expanded to include families from Orange County enrolled by the University of California, Irvine.

Protocols and procedures

Initially six Working Groups were established to develop uniform procedures and questionnaires for data and biospecimen collection and processing: Family History, Epidemiology, Biospecimens, Pathology, Database, and Informed Consent. These Working Groups developed the instruments for data collection, the protocols for biospecimen collection, processing, and distribution, and the data dictionaries to be used at the Informatics Center. An early challenge was to recognize and respect the geographic differences in social, cultural and health care structures and legislation, while finding common principles, issues, and language for the questionnaires and informed consent forms. Core epidemiology and treatment questionnaires were developed, and common language was incorporated into each site-specific consent form to address issues faced by all sites [14].

All sites collected the following: family history data from probands; epidemiological and dietary data, and blood samples (or mouthwash samples if venipuncture was declined) from probands and selected relatives; and clinical and treatment data, tumor blocks and pathology reports for probands and relatives with a personal history of breast or ovarian cancer. All data and biospecimens were stored without personal identifiers.

Questionnaires

Family history questionnaire

Information was sought, at minimum, about previous cancer diagnoses in the proband and the proband's parents, siblings, and children. Similar information for more distant relatives was also sought depending on site protocols. All cancers, except non-melanoma skin cancers and cervical carcinoma *in situ*, were recorded. Dates of all cancer diagnoses and deaths were requested.

Epidemiology questionnaire

This instrument obtained information on demographics, race/ethnicity, religion, personal history of cancer, breast

and ovarian surgeries, radiation exposure, smoking and alcohol consumption, menstrual and pregnancy history, breast-feeding, hormone use, weight, height, and physical activity. Some sites used a short proxy version to collect limited information on deceased relatives and selected living relatives.

Dietary questionnaires

A self-administered food frequency questionnaire developed by the University of Hawaii [15] for multi-ethnic cohort studies was used by the five North American sites. The Melbourne and Sydney site used a locally validated dietary questionnaire developed for a cohort study of Greek, Italian, and Australian-born inhabitants of Melbourne [16]. Both instruments collected information on frequency of food consumption and portion size, using photographs to help in assigning portion sizes.

Treatment questionnaire

Self-reported information was sought on aspects of treatment for breast or ovarian cancer and for any recurrences. Information from medical records was also collected by some sites.

Biospecimen collection and processing

A 30 ml sample of blood was requested from probands and selected relatives, and paraffin blocks or unstained sections of the paraffin blocks were requested for individuals with a history of breast or ovarian cancer. From participants who declined venipuncture, a mouthwash sample was collected at some sites in accordance with the protocol of Lum and Le Marchand [17]. The New York site also collected urine samples from selected participants for estrogen metabolite analyses.

Blood and mouthwash samples

Biospecimen samples were processed at each site, or at a collaborating laboratory, in accordance with a common standardized protocol. A quality control program was developed to allow validation of the methods and their application at each site. From three tubes of blood, one tube was used for direct DNA isolation, a second was used for the preparation of blood spots and plasma, and a third was collected for the isolation and cryopreservation of lymphocytes for future transformation or DNA preparation. To provide an unlimited source for nucleic acids, Epstein-Barr virus-transformed lymphoblastoid cell lines from probands and selected relatives were established [18]. Biospecimens were stored at either the participating academic institutions or, for some sites, at the Coriell Institute for Medical Research. Biospecimen collection, processing, annotation, storage, and distribution for the Breast Cancer Family Registry were evaluated recently in a report prepared for the National Cancer Institute and the National Dialogue on Cancer [19] and it was noted that many of the 'best

practices' suggested for biospecimen repositories are currently in use within the Breast Cancer Family Registry.

Pathology specimens

For individuals with a personal history of breast or ovarian cancer, histological slides and/or paraffin tumor blocks were requested from the treating institution. Sections were cut from each block, stained with hematoxylin and eosin, and reviewed by the site pathologist(s). The pathologists used standard pathology review forms that were developed by the Pathology Working Group. A set of 35 invasive carcinomas were reviewed by pathologists at all sites, and agreement was found to be good to excellent for several pathologic characteristics (Longacre TA, Bane A, Bleiweiss I, Carter B, Catelano E, Ennis M, Hendrickson MR, Hibshoosh H, Layfield L, Memeo L, Quenneville L, Venter DJ, Wu H, O'Malley FP, unpublished data), thus validating the use of the semi-centralized review process instituted within the Breast Cancer Family Registry. Representative blocks of tumors and also of associated benign lesions were selected by the pathologists for retention in the tissue repository. All other slides and blocks were returned to the treating institution. If permission was not obtained to retain the representative blocks in the repository, sections were cut in accordance with a standard cutting protocol. This included cutting 10–20 sections at 4 μm thickness for future immunohistochemical studies and an additional 10–20 sections at 10 μm for future DNA extraction. Control sections, for staining with hematoxylin and eosin, were taken at the beginning, middle and end of the cutting protocol as a quality control measure. The specific number of sections taken from a block depended on the amount of tumor present in the block. The slides sectioned for future immunohistochemical studies were placed in either +4°C fridges or -20 or -80°C freezers. This was to minimize the risk of any loss of antigenicity, which is known to occur if unstained sections are stored at room temperature. If permission had been obtained to retain tumor blocks in the repository, further permission was sought to construct tissue microarrays from these blocks. Tissue microarrays have been constructed at the Ontario site and are soon to be constructed at other sites. All pathology reviews were entered into a database and submitted to the Informatics Center at the University of California at Irvine.

Validation of breast and ovarian cancer diagnoses

Verification was sought for all reported breast and ovarian cancers, and at some sites for all reported cancers. Because the population-based sites ascertained case probands through cancer registries, verification was necessary only for cancers reported for relatives. The level of confidence regarding a cancer diagnosis was classified into one of six categories, in decreasing order: (1) review of slides by Breast Cancer Family Registry pathologist, (2) pathology report, (3) cancer registry report or medical

records indicating treatment for the specific type of cancer, (4) report on a death certificate, (5) self-report, and (6) report by a relative.

Follow-up

Selected participants are being followed to obtain updated information on cancer and vital status of family members. For the clinic-based families in the USA, at least one participant from each family is contacted annually to update personal and family cancer histories and deaths, as well as some exposures addressed in the core epidemiology questionnaire. In Ontario, an annual mailed follow-up questionnaire to case probands seeks to update births, deaths, and new cancer diagnoses of case probands and family members. The San Francisco site contacts case probands annually by telephone to update information on cancer and vital status of the proband and family members. In Australia, passive record linking to state cancer registries and death certificates is being conducted, and the medical records of case probands have been followed up for recurrence and death [20]. A systematic registry-wide follow-up of all enrolled probands and relatives is being developed by the Follow-up Working Group, and is currently being pilot tested in Australia, Ontario, and San Francisco.

Mutational analysis of BRCA1 and BRCA2

Substantial mutational analyses of *BRCA1* and *BRCA2* have been undertaken by site laboratories and, more recently, by Myriad Genetics using full sequence analysis [21], funded by multiple sources. A validation study was conducted for five of the methods used between 1997 and 2000, including four DNA-based methods (namely two-dimensional gene scanning, denaturing high-performance liquid chromatography, enzymatic mutation detection, and single-strand conformation polymorphism analysis) and an RNA/DNA-based method (a protein truncation test) [22]. Single-strand conformation polymorphism analysis was less sensitive than the other methods and is no longer being used. The specificity and sensitivity of the other four methods for protein-truncating mutations were comparable to those of full sequencing ('gold standard'). Ashkenazi Jewish participants have been screened for the three founder mutations, 185delAG and 5382insC in *BRCA1* and 6174delT in *BRCA2*.

Results

Collection of data and biospecimens

As of September 2003, the six sites had enrolled a total of 11,950 families in the Breast Cancer Family Registry (Table 3). They included 6126 population-based case families and 2990 population-based control families, and 1647 clinic-based families with an affected proband and 1187 clinic-based families with an unaffected proband. Affected probands included 7111 females with a first primary breast cancer, 538 females with a second breast cancer, and 124

Table 3**Breast cancer status and sex of probands, by ascertainment source and recruitment site (1996–2003)**

Source of proband ascertainment	San Francisco ^a	Ontario	Melbourne and Sydney	New York	Philadelphia	Utah	Total
Population-based case probands							6126
Female, with first primary breast cancer	2038	2553	1141				5732
Female, with second breast cancer	199	91	0				290
Male	56	48	0				104
Population-based control probands							2990
Female	634	1711	645				2990
Clinic-based affected probands							1647
Female, with first primary breast cancer	0	333	177	574	241	54	1379
Female, with second breast cancer	0	27	15	152	34	20	248
Male	0	3	1	14	2	0	20
Clinic-based unaffected probands							1187
Female	0	0	179	386	489	109	1163
Male	0	0	2	8	14	0	24
Total probands	2927	4766	2160	1134	780	183	11,950
Female, with first primary breast cancer	2038	2886	1318	574	241	54	7111
Female, with second breast cancer	199	118	15	152	34	20	538
Female, without breast cancer	634	1711	824	386	489	109	4153
Male, with breast cancer	56	51	1	14	2	0	124
Male, without breast cancer	0	0	2	8	14	0	24

^aIncludes probands recruited in Orange County, California, since 2001.

males with breast cancer. The enrolled families included 2346 minority families, residing mostly in the USA.

The epidemiology questionnaire was completed by 27,421 participants (10,895 probands and 16,526 relatives), and the short proxy epidemiology questionnaire was completed for 20,003 relatives (Table 4). Blood or mouthwash samples were collected from 20,045 individuals (9282 probands and 10,763 relatives), and tumor tissue was obtained for 4293 individuals with a history of breast and/or ovarian cancer (3322 probands and 971 relatives).

Population-based families

Because recruitment of families is continuing in San Francisco and Ontario, the participation rates below refer to enrollment from 1996–2000. At each site, before contact was made with incident breast cancer cases identified from the regional cancer registry, the case's physician was contacted. Physician consent was obtained to contact the great majority of case probands (98% in San Francisco, 92% in Ontario, and 90% in Melbourne and Sydney). In Ontario and in Melbourne and Sydney, 2% and 3%, respectively, of the case probands were deceased and

were therefore not studied at those sites. In San Francisco, family history and epidemiology data for deceased case probands were collected from proxy respondents.

Eligibility of case probands

To determine eligibility for sampling as a case proband, information on family history of breast or ovarian cancer was first obtained through a telephone interview in San Francisco (84% response rate), and by a mailed questionnaire in Ontario (65% response rate). In Melbourne and Sydney, all newly diagnosed breast cancer cases were eligible, regardless of family history of breast cancer.

Case probands

Of the eligible case probands, 6126 completed the family history questionnaire, including 104 males (Table 3), and 5250 completed the epidemiology and treatment questionnaires (76% in San Francisco, 72% in Ontario, and 75% in Melbourne and Sydney for both the family history and epidemiology questionnaires) (Table 4). An analysis at the Melbourne and Sydney site showed that there was high agreement between self-reported treatment data and medical records (Phillips KA, Milne RL, Buys S, Friedlander ML,

Table 4**Epidemiology data and biospecimen collection for probands and relatives from all sites (1996–2003), by ascertainment source**

Parameter	Population-based families		Clinic-based families		Total families
	Probands	Relatives	Probands	Relatives	
Affected probands and relatives ^a	6126	22,857	1647	4943	35,573
Affected (breast only)	6098	1856	1602	694	10,250
Affected (ovarian only)	0	158	0	81	239
Affected (breast and ovarian)	28	24	45	25	122
With epidemiology questionnaire	5250	10,535	1529	2857	20,171
With short proxy epidemiology questionnaire	0	11,155	0	1821	12,976
With blood sample	4429	6293	1524	2204	14,450
With mouthwash sample	357	483	6	31	877
With tumor tissue ^b	2675	437	630	151	3893
Unaffected probands and relatives ^a	2990	7229	1187	3321	14,727
With epidemiology questionnaire	2979	1387	1137	1747	7250
With short proxy epidemiology questionnaire	0	5842	0	1185	7027
With blood sample	1797	14	1108	1731	4650
With mouthwash sample	58	0	3	7	68
With tumor tissue ^b	0	1	17	382	400
All probands and relatives ^a	9116	30,086	2834	8264	50,300
With epidemiology questionnaire	8229	11,922	2666	4604	27,421
With short proxy epidemiology questionnaire	0	16,997	0	3006	20,003
With blood sample	6226	6307	2632	3935	19,100
With mouthwash sample	415	483	9	38	945
With tumor tissue ^b	2675	438	647	533	4293

^aProbands with completed family history questionnaire; relatives with epidemiology data, blood or mouthwash sample, or tumor block. ^bTumor tissue for breast or ovarian cancer.

Ward J, McCredie MRE, Giles GG, Hopper JL, unpublished data).

Blood or mouthwash samples were collected from 4786 case probands (70% in San Francisco, 62% in Ontario, and 71% in Melbourne and Sydney) (Table 4). An analysis of participants at the Ontario site showed that proband non-response at all stages (namely family history, epidemiology questionnaire, and biospecimen collection) was not associated with family history of breast or ovarian cancer [23,24]. Lymphoblastoid cell lines were established for 1723 case probands, and tumor tissue samples were obtained for 2675.

Relatives

A total of 22,857 relatives of case probands have been enrolled (Table 4). The epidemiology questionnaire was

completed by 10,535 relatives. The short proxy epidemiology questionnaire was completed for 11,155 relatives from the Australian site. Blood or mouthwash samples were collected for 6776 relatives. Lymphoblastoid cell lines were established for 450 relatives, and tumor tissue was obtained for 437 affected relatives. Collection of tumor tissue is still continuing in Ontario and in Melbourne and Sydney.

Control probands

Among women selected as control probands, 2990 completed the family history questionnaire, 2979 completed the epidemiology questionnaire, and 1855 provided a blood or mouthwash sample. Response to the epidemiology questionnaire was 60% in San Francisco, 64% in Ontario, and 68% in Melbourne and Sydney. Participation in biospecimen collection was 56% in San Francisco and

55% in Melbourne and Sydney. Biospecimen collection is continuing in Ontario and is expected to be completed for 45% of control probands.

The same ascertainment protocol was also used in Melbourne and Sydney to recruit case and control probands before the establishment of the Breast Cancer Family Registry, including 467 case probands diagnosed between 1992 and 1995 with breast cancer before the age of 40 years, and 408 control probands frequency-matched to case probands on age. Relevant data are stored at the Informatics Center and, together with material collected from family members, are available from the site investigators to be used in conjunction with the Breast Cancer Family Registry resources.

Clinic-based families

A total of 2834 probands (1647 affected, 1187 unaffected) have been enrolled (Table 4). Of these, 2666 completed the epidemiology questionnaire and 2641 provided a blood or mouthwash sample (1530 affected, 1111 unaffected). Lymphoblastoid cell lines were established for 738 (343 affected, 395 unaffected). Tumor blocks were obtained for 647 probands.

A total of 8264 relatives have been enrolled, with an average of three members per family. Of these, 4604 completed the epidemiology questionnaire, and 3973 provided a blood or mouthwash sample. The short proxy epidemiology questionnaire was completed for 3006 relatives. Lymphoblastoid cell lines were established for 1014 relatives. Tumor blocks for breast or ovarian cancer were obtained for 533 relatives.

In addition to families presented in Table 4, more than 500 multiple-case breast cancer families have been recruited in Australia as part of the Kathleen Cuninghame Consortium for Familial Breast Cancer (kConFab), which administered the same epidemiology and family history questionnaires and used the same blood collection protocol as the Breast Cancer Family Registry. Funds for *BRCA1* and *BRCA2* mutation testing have been provided by the National Cancer Institute and these families are available to be used in conjunction with the Breast Cancer Family Registry resources through application to kConFab <http://www.kconfab.org>.

Proband and family characteristics

The Breast Cancer Family Registry contains various subgroups of probands and families with specific characteristics (Table 5). Among the 6779 probands with a history of breast or ovarian cancer and a completed epidemiology questionnaire (5250 from population-based families and 1529 from clinic-based families) there are 124 male probands, 1526 (23%) with a diagnosis before age 40

years, 1748 (26%) from minority populations, 1040 (15%) of Ashkenazi Jewish ancestry, 494 (7%) with a history of two breast cancer diagnoses, 65 (1%) with a history of both breast and ovarian cancer, 2332 (34%) with at least one first-degree relative with breast cancer, and 61 from participating twin pairs. The relatively high proportion (23%) of probands diagnosed before the age of 40 years reflects both the designs used by the population-based sampling to increase the number of case probands with a genetic etiology and the age at diagnosis distribution of the multiple-case families. The relatively large proportion (26%) of minority probands largely reflects the oversampling of these families at the San Francisco site. Among them, 25% are Latino, 24% are African-American, 15% are Chinese, 13% are Filipino, 4% are Japanese, and the remaining 19% are other Asians, Pacific Islanders and others. Probands reported 101 different countries of birth; 60% were born in the USA or Canada, 14% in Australia or New Zealand, 12% in Europe, 9% in Asia, 4% in Latin America, and 1% in Africa.

Among the 2834 clinic-based families enrolled so far, 204 (7%) include three or more first-degree relatives with breast or ovarian cancer (Table 6). Among the 6126 population-based case families this percentage is 4%. Blood or mouthwash samples are available for 1863 sibships from population-based case families and 701 sibships from clinic-based families with one or more affected sisters and one or more unaffected sisters (Table 7).

Mutational analysis of *BRCA1* and *BRCA2*

Testing for mutations in *BRCA1* has been conducted for 5656 females and 612 males, and in *BRCA2* for 5497 females and 524 males (Table 8). Nearly half of those tested (42%) were of Ashkenazi Jewish ancestry. A total of 984 mutation carriers have been detected (547 affected with breast cancer, 437 unaffected). Among 539 female *BRCA1* mutation carriers detected, 329 (61%) had a history of breast cancer, in comparison with 207 (70%) among 297 female *BRCA2* mutation carriers detected. There are a total of 230 population-based female mutation carriers, making this the largest collection of population-based carriers yet established. At some sites, *BRCA1* and *BRCA2* mutation testing is continuing; the number of available mutation carriers will therefore increase.

Discussion

The Breast Cancer Family Registry has enrolled nearly 12,000 families containing individuals with a wide range of familial risks of breast cancer. This novel research infrastructure has many strengths, including the following: its focus on data and biospecimen collection from both population-based and clinic-based families, including a large number of minority families and Ashkenazi Jewish families from three countries; attention to quality, compar-

Table 5**Age and race/ethnicity of probands with a history of breast cancer from all sites (1996–2003), by ascertainment source**

Parameter	Population-based case probands ^b (<i>n</i> = 5250)	Clinic-based affected probands ^b (<i>n</i> = 1529)	Total probands ^b (<i>n</i> = 6779)
Age at diagnosis			
<30 years	132	45	177
30–39 years	1049	300	1349
40–49 years	1809	585	2394
50–59 years	1610	344	1954
≥ 60 years	650	255	905
Race/ethnicity			
White	3604	1278	4882
African-American	381	45	426
Latino	330	105	435
Chinese	248	21	269
Japanese	66	0	66
Filipino	227	0	227
Other Asian or Pacific Islander	122	7	129
Mixed or other race	130	66	196
Not reported	142	7	149
Ashkenazi Jewish heritage	232	808	1040
History of second breast cancer	261	233	494
History of breast and ovarian cancer	22	43	65
Family history of breast cancer in one or more first-degree relatives	1677	655	2332

^aProbands with completed epidemiology questionnaire. ^bProbands with a history of breast cancer.

Table 6**Distribution of families from all sites (1996–2003), by history of breast and ovarian cancer and ascertainment source**

Number of reported breast and ovarian cancers in family ^a	Number of families enrolled		
	Population-based case families (<i>n</i> = 6126)	Clinic-based families (<i>n</i> = 2834)	Total families (<i>n</i> = 8960)
0	0 (0%)	152 ^b (5%)	152 (1%)
1	3979 (65%)	1684 (59%)	5663 (63%)
2	1873 (30%)	794 (28%)	2667 (30%)
3	240 (4%)	178 (6%)	418 (5%)
4	24 (<1%)	19 (<1%)	43 (<1%)
≥ 5	10 (<1%)	7 (<1%)	17 (<1%)

^aIn proband and first-degree relatives. ^bBreast or ovarian cancers in second-degree relatives only.

bility, and comprehensiveness of data and biospecimen collection; continuing molecular characterization; establishment of Epstein–Barr virus-immortalized lymphoblastoid

cell lines; and broad-based research and clinical expertise in breast cancer represented among the six participating sites and their international collaborators.

Table 7**Distribution of sibships from all sites (1996–2003), by number of affected and unaffected sisters and ascertainment source**

Number of affected sisters ^a	Number of unaffected sisters ^a	Number of sibships		
		Population-based case families	Clinic-based families	Total families
1	0	4130	1649	5779
1	1	979	377	1356
1	2	382	105	487
1	≥ 3	157	28	185
2	0	169	118	287
2	1	83	42	125
2	2	46	9	55
2	≥ 3	15	2	17
≥ 3	0	20	14	34
≥ 3	1	8	2	10
≥ 3	2	0	4	4
≥ 3	≥ 3	4	0	4
Total sibships		5993	2350	8343

^aFull sisters with blood or mouthwash samples.

The Breast Cancer Family Registry includes families identified through multiple ascertainment modes, thus permitting flexibility in the design of studies using the resources. The infrastructure fosters collaborative interdisciplinary studies to rapidly address a broad range of research questions requiring large samples of data and biospecimens that are readily available and well defined in terms of epidemiological, clinical, and molecular characteristics. Thus, the Breast Cancer Family Registry provides a model for studying the genetic epidemiology of other cancers and other complex diseases.

Extended families with multiple cases are a proven means of discovering genes that when mutated convey a high risk of disease [25], and are an efficient sampling design for identifying mutation carriers for studies of genetic and environmental risk modifiers. Sisters concordant for disease are useful for gene discovery and for disentangling gene–gene interactions, and sisters discordant for disease are useful for case-control studies of putative genetic and environmental risk factors. Population-based case families can be used to characterize susceptibility genes by providing estimates of penetrance and prevalence applicable to the population groups from which they are sampled, and, when combined with controls and control families, can provide multiple designs for addressing issues in the genetic epidemiology of breast cancer [26–29]. The availability of population-based and family-based controls within the registry, combined with the ethnic diversity of the families, will per-

mit questions related to population stratification to be addressed. Populations carrying founder or ancestral mutations in susceptibility genes can facilitate the characterization of specific mutations and their impact on communities, whereas twin studies can provide new insights into the relative roles of genetic and environmental factors [30,31]. Population-based case probands and both related and unrelated controls are important for assessing the effects of measured genetic variants or haplotypes in candidate genes [32], which might have a low individual risk but, when combined, the genetic variants and haplotypes might have a high population attributable risk, with substantial relevance to public health.

The different designs and common resources available through the Breast Cancer Family Registry can be used for a multitude of collaborative studies; see, for example, Whittemore and Nelson [33] for a discussion of designs, strengths, and weaknesses. Some of the studies and initiatives currently under way or in development include searching for novel breast cancer susceptibility loci, testing for association and/or linkage with variants in known candidate genes, estimating the penetrance and detecting modifiers of penetrance associated with variants in different genes (including the examination of genetic and environmental modifiers of risks associated with these variants, often referred to as gene–gene and gene–environment interactions), and developing and disseminating innovative analytical approaches and related software for the discovery and

Table 8**Number of currently identified BRCA1 or BRCA2 mutation carriers^a from all sites (1996–2003), by sex and ascertainment source**

Sex and group	Status	Population-based case families	Clinic-based families	Families from Ashkenazi Jewish Project	Total families
Females					
Tested for <i>BRCA1</i>		2252	1029	2375	5656
<i>BRCA1</i> carriers	Affected ^b	110	70	149	329
	Unaffected	23	100	87	210
	Total	133	170	236	539
Tested for <i>BRCA2</i>		2255	868	2374	5497
<i>BRCA2</i> carriers	Affected ^b	85	59	63	207
	Unaffected	12	36	42	90
	Total	97	95	105	297
Males					
Tested for <i>BRCA1</i>		86	137	389	612
<i>BRCA1</i> carriers	Affected ^b	0	1	3	4
	Unaffected	14	44	41	99
	Total	14	45	44	103
Tested for <i>BRCA2</i>		66	73	385	524
<i>BRCA2</i> carriers	Affected ^b	3	3	1	7
	Unaffected	7	13	18	38
	Total	10	16	19	45
Males and females					
<i>BRCA1</i> or <i>BRCA2</i> carriers	Affected	198	133	216	547
	Unaffected	56	193	188	437
	Total	254	326	404	984

^aCarriers of deleterious mutations. ^bWith a history of breast cancer.

characterization of cancer susceptibility genes. In addition, clinical epidemiology studies are addressing prognosis and optimal treatment for various high-risk subgroups. Social and behavioral epidemiology studies are addressing choices by, and behaviors of, high-risk individuals. Future health policy and public health research might focus on interrelationships between different legislative, health care, and social structures and health policies related to genetic testing and research.

Potential limitations of the Breast Cancer Family Registry need to be considered when using its resources. There are differences across and even within sites in designs, eligibility criteria, sampling schemes, and data collection modes that can be used to advantage, but can be problematic if not understood. Proper analysis and interpretation of family studies across sites require a clear understanding of the ascertainment procedures that have been used. As in all

family studies, biospecimens are not available from all eligible family members, and study designs and analyses need to consider this limitation. Lastly, there is also some incompleteness in the collection of questionnaire data and tumor samples. The impact of each of these issues on specific studies needs to be considered and, if possible, minimized.

The Breast Cancer Family Registry offers several challenges for theoretical and applied statisticians in developing optimal methods for the design and analysis of studies using its resources. Such challenges include the following: analyzing data from individuals who are related and from families for whom data collection is incomplete; trying to make inferences about either a measured genetic marker or about the characteristics (mode of inheritance, allele frequency, effects on risk) of a presumed unmeasured genetic effect against the background of other familial effects of unknown origin, such as polygenic inheritance or shared

family environmental factors [32,34]; and how to make appropriate adjustment for non-random or non-systematic ascertainment of families. Researchers from the Breast Cancer Family Registry are collaborating with other statisticians to develop methods and software to facilitate appropriate and optimal analyses and to make these developments available to researchers using the resource.

The development of the Breast Cancer Family Registry has already resulted in the initiation of more than 80 hypothesis-driven research projects among participating sites and with the greater international research community, and has already produced numerous publications. The Breast Cancer Family Registry website lists continuing collaborative research projects http://www.cfr.epi.uci.edu/ic_registries/breast/approved.htm and publications http://www.cfr.epi.uci.edu/ic_registries/breast/breast_publications.htm.

The Breast Cancer Family Registry data and biospecimens are available to the scientific community. Researchers interested in initiating collaborative research projects using the resources are invited to access the Breast Cancer Family Registry website for preliminary information <http://epi.grants.cancer.gov/BCFR/index.html>, and to make initial contact with the Program Officer at the National Cancer Institute to discuss the process of developing a collaborative proposal (details are available from the corresponding author). Interested investigators will then be referred to the relevant investigators and Working Groups to discuss the details of their proposal, to become acquainted with site-specific recruitment issues, to establish collaborations, and then to submit to the Advisory and Steering Committees a concise proposal that includes information on the study design and requirements for data and/or biospecimens. Approval for the use of human subjects in accordance with the requirements of the Office for Human Research Protections is required. For approved proposals requesting the use of biospecimens, a Material Transfer Agreement is required.

Conclusion

Nearly 12,000 families have been enrolled in the Breast Cancer Family Registry, a novel research infrastructure. Data and biospecimen resources are available for collaborative, interdisciplinary, and translational studies of the genetic epidemiology of breast cancer.

Competing interests

None declared.

Acknowledgements

This study was funded by the United States National Cancer Institute, National Institutes of Health, under Request for Application CA-95-003 as part of the Breast Cancer Family Registries, and through cooperative agreements with the Fox Chase Cancer Center; Huntsman Cancer Insti-

tute; Columbia University; the Northern California Cancer Center; Cancer Care Ontario; The University of Melbourne, and The University of California Irvine. It was also supported in the USA by NIH grants U01CA 71966, CA13696, and ES09089, a Public Health Service research grant (MO1-RR00064) from the National Center for Research Resources, and the Huntsman Cancer Foundation, in Canada by Cancer Care Ontario and the Canadian Breast Cancer Research Alliance, and in Australia by the Australian National Health and Medical Research Council. The contents of the article are solely the responsibility of the authors and do not necessarily represent the official views of the organizations named above. We wish to express our appreciation to the women and men who participated in this study, and to the medical practitioners who supported the concept.

References

1. Peto J: **Genetic predisposition to cancer.** In *Cancer Incidence in Defined Populations* Edited by: Cairns J, Lyon JL, Skolnick M. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory; 1980:203-213. Banbury Report no. 4.
2. Hopper JL, Carlin JB: **Familial aggregation of a disease consequent upon correlation between relatives in a risk factor measured on a continuous scale.** *Am J Epidemiol* 1992, **136**:1138-1147.
3. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, Bell R, Rosenthal J, Hussey C, Tran T, McClure M, Frye C, Hattier T, Phelps R, Haugen-Strano A, Katcher H, Yakumo K, Gholami Z, Shaffer D, Stone S, Bayer S, Wray C, Bogden R, Dayananth P, Ward J, Tonin P, Narod S, Bristow PK, Norris FH, Helvering L, Morrison P, Rostek P, Lai M, Barrett JC, Lewis C, Neuhausen S, Cannon-Albright L, Goldgar D, Wiseman R, Kamb A, Skolnick MH: **A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.** *Science* 1994, **266**:66-71.
4. Wooster R, Neuhausen S, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, Seal S, Tran T, Averill D, Fields P, Marshall G, Narod S, Lenoir G, Lynch H, Devilee P, Cornelisse CJ, Menko FH, Daly PA, Ormiston W, McManus R, Pye C, Cannon-Albright L, Peto J, Ponder BAJ, Skolnick MH, Easton DF, Goldgar DE, Stratton MR: **Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.** *Science* 1994, **265**:2088-2090.
5. Peto J, Collins N, Barfoot R, Seal S, Warren W, Rahman N, Easton DF, Evans C, Deacon J, Stratton MR: **Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer.** *J Natl Cancer Inst* 1999, **91**:943-949.
6. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J, Sobol H, Teare MD, Struwing J, Arason A, Scherneck S, Peto J, Rebbeck TR, Tonin P, Neuhausen S, Barkardottir R, Eyfjord J, Lynch H, Ponder BA, Gayther SA, Birch MJ, Lindblom A, Stoppa-Lyonnet D, Bignon Y, Borg A, Hamann U, Haites N, Scott RJ, Maugard CM, Vasen H, Seitz S, Cannon-Albright LA, Schofield A, Zelada-Hedman M, the Breast Cancer Linkage Consortium: **Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families.** *Am J Hum Genet* 1998, **62**:676-689.
7. Børresen AL, Andersen TI, Garber J, Barbier-Pirax N, Thorlacius S, Eyfjord J, Ottstad L, Smith-Sørensen B, Hovig E, Malkin D: **Screening for germ line TP53 mutations in breast cancer patients.** *Cancer Res* 1992, **52**:3234-3236.
8. Chenevix-Trench G, Spurdle AB, Gatei M, Kelly H, Marsh A, Chen X, Donn K, Cummings M, Nyholt D, Jenkins MA, Scott C, Pupo GM, Dork T, Bendix R, Kirk J, Tucker K, McCredie MRE, Hopper JL, Sambrook J, Mann GJ, Khanna KK: **Dominant negative ATM mutations in breast cancer families.** *J Natl Cancer Inst* 2002, **94**:205-215.
9. Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, Oldenburg R, Hollestelle A, Houben M, Crepin E, van Veghel-Plandsoen M, Elstrodt F, van Duijn C, Bartels C, Meijers C, Schutte M, McGuffog L, Thompson D, Easton D, Sodha N, Seal S, Barfoot R, Mangion J, Chang-Claude J, Eccles D, Eeles R, Evans DG, Houlston R, Munday V, Narod S, Peretz T, Peto J, Phelan C, Zhang HX, Szabo C, Devilee P, Goldgar D, Futreal PA, Nathanson KL, Weber B, Rahman N, Stratton MR: **Low-penetrance suscep-**

- ability to breast cancer due to CHEK2*1100delC in noncarriers of *BRCA1* or *BRCA2* mutations. *Nat Genet* 2002, **31**:55-59.
10. Grabrick DM, Hartmann LC, Cerhan JR, Vierkant RA, Therneau TM, Vachon CM, Olson JE, Couch FJ, Anderson KE, Pankratz VS, Sellers TA: **Risk of breast cancer with oral contraceptive use in women with a family history of breast cancer.** *JAMA* 2000, **284**:1791-1798.
 11. Jernstrom H, Lerman C, Ghadirian P, Lynch HT, Weber B, Garber J, Daly M, Olopade OI, Foulkes WD, Warner E, Brunet JS, Narod SA: **Pregnancy and risk of early breast cancer in carriers of *BRCA1* and *BRCA2*.** *Lancet* 1999, **354**:1846-1850.
 12. Greenspan RJ: **The flexible genome.** *Nat Rev Genet* 2001, **2**:383-387.
 13. Roa BB, Boyd AA, Volcik K, Richards CS: **Ashkenazi Jewish population frequencies for common mutations in *BRCA1* and *BRCA2*.** *Nat Genet* 1996, **14**:185-187.
 14. Daly MB, Offit K, Li F, Glendon G, Yaker A, West D, Koenig B, McCredie M, Venne V, Nayfield S, Seminara D: **Participation in the Cooperative Family Registry for Breast Cancer Studies (CFRBCS): issues of informed consent.** *J Natl Cancer Inst* 2000, **92**:452-456.
 15. Hankin JH, Wilkens LR, Kolonel LN, Yoshizawa CN: **Validation of a quantitative diet history method in Hawaii.** *Am J Epidemiol* 1991, **133**:616-628.
 16. Ireland P, Jolley D, Giles GG: **Development of the Melbourne FFO: a food frequency questionnaire for use in an Australian prospective study involving an ethnically diverse cohort.** *Asia Pac J Clin Nutr* 1994, **3**:19-31.
 17. Lum A, Le Marchand L: **A simple mouthwash method for obtaining genomic DNA in molecular epidemiologic studies.** *Cancer Epidemiol Biomarkers Prev* 1998, **7**:719-724.
 18. Beck JC, Beiswanger CM, John EM, Satariano E, West D: **Successful transformation of cryopreserved lymphocytes: a resource for epidemiological studies.** *Cancer Epidemiol Biomarkers Prev* 2001, **10**:551-554.
 19. Eiseman E, Bloom G, Brower J, Clancy N, Olmsted S: *Case Studies of Existing Human Repositories: Best Practices for a Biospecimen Resource for the Genomic and Proteomic Era* Santa Monica, CA: Rand Corporations Publishers; 2003.
 20. Phillips KA, Milne RL, Friedlander ML, Jenkins MA, McCredie MRE, Giles GG, Hopper JL: **Prognosis of premenopausal breast cancer and childbirth prior to diagnosis.** *J Clin Oncol* 2004, **22**:699-705.
 21. Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B, Gumpfer KL, Scholl T, Tavtigian SV, Pruss DR, Critchfield GC: **Clinical characteristics of individuals with germline mutations in *BRCA1* and *BRCA2*: analysis of 10,000 individuals.** *J Clin Oncol* 2002, **20**:1480-1490.
 22. Andrulis IL, Anton-Culver H, Beck J, Bove B, Boyd J, Buys S, Godwin A, Hopper J, Li F, Neuhausen SL, Ozcelik H, Peel D, Santella RM, Southey M, Van Orsouw NJ, Venter D, Vijg J, Whittemore AS, for the Cooperative Family Registry for Breast Cancer Studies: **Comparison of DNA- and RNA-based methods for detection of truncating *BRCA1* mutations.** *Hum Mutat* 2002, **20**:65-73.
 23. Knight JA, Sutherland HJ, Glendon G, Boyd NF, Andrulis IL, the Ontario Cancer Genetics Network: **Characteristics associated with participation at various stages at the Ontario site of the Cooperative Family Registry for Breast Cancer Studies.** *Ann Epidemiol* 2002, **12**:27-33.
 24. Mancuso C, Glendon G, Anson-Cartwright L, Shi EJO, Andrulis IL, Knight JA: **Ethnicity, but not cancer family history, is related to response to a population-based mailed questionnaire.** *Ann Epidemiol* 2004, **14**:36-43.
 25. Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M: **Genomewide scans of complex human diseases: true linkage is hard to find.** *Am J Hum Genet* 2001, **69**:936-950.
 26. Hopper JL: **Commentary: Case-control-family designs: a paradigm for future epidemiology research?** *Int J Epidemiol* 2003, **32**:48-50.
 27. Thomas DC, Witte JS: **Point: population stratification: a problem for case-control studies of candidate-gene associations?** *Cancer Epidemiol Biomarkers Prev* 2002, **11**:505-512.
 28. Whittemore AS, Halpern J: **Multi-stage sampling in genetic epidemiology.** *Stat Med* 1997, **16**:153-167.
 29. Whittemore AS, Halpern J: **Logistic regression of family data from retrospective study designs.** *Genet Epidemiol* 2003, **25**:177-189.
 30. Peto J, Mack TM: **High constant incidence in twins and other relatives of women with breast cancer.** *Nat Genet* 2000, **26**:411-414.
 31. Mack TM, Hamilton AS, Press MF, Diep A, Rappaport EB: **Heritable breast cancer in twins.** *Br J Cancer* 2002, **87**:294-300.
 32. Cui JS, Spurdle AB, Southey MC, Dite GS, Venter DJ, McCredie MR, Giles GG, Chenevix-Trench G, Hopper JL: **Regressive logistic and proportional hazards disease models for within-family analyses of measured genotypes, with application to a CYP17 polymorphism and breast cancer.** *Genet Epidemiol* 2003, **24**:161-72.
 33. Whittemore AS, Nelson LM: **Study design in genetic epidemiology: theoretical and practical considerations.** *J Natl Cancer Inst Monogr* 1999, **26**:61-9.
 34. Antoniou A, Pharoah PD, McMullan G, Day NE, Stratton MR, Peto J, Ponder BJ, Easton DF: **A comprehensive model for familial breast cancer incorporating *BRCA1*, *BRCA2* and other genes.** *Br J Cancer* 2002, **86**:76-83.