

## **BREAST CANCER FAMILY REGISTRY (BCFR) DATA SHARING PLAN**

**What data will be shared:** All questionnaire data, clinical data, and biospecimen data collected by the BCFR will be shared after approval has been granted to use our Resource (see below).

**Who will have access to the data:** Any investigator with an approved project and demonstrated Data Use agreements and/or Material Transfer Agreements and IRB approval will have access to the data.

**Where will the data to be shared be located:** All non-biospecimen data are stored centrally at the site of the BCFR central data core, also known as the Informatics Support Center (ISC), a comprehensive central database to process and store data at Columbia University. Biospecimens are stored locally at each of the six sites of the BCFR.

**When will the data be shared:** The data will be shared as soon as the project is approved, IRB is approved, and the Data Use agreement is signed by both Columbia University and the receiving institution.

**How will researchers locate and access the data:** Upon approvals, researchers will access data through our secured server at Columbia University.

**Additional details:** In addition to publishing and presenting our findings at international, national and local conferences and meetings, our primary mechanism for sharing BCFR resources (data and/or biospecimens) has been and will continue to conduct collaborative multinational, interdisciplinary and translational studies organized through the BCFR. The data and/or biospecimens that are available from the BCFR include but are not limited to detailed epidemiologic questionnaire data (e.g., lifestyle and dietary factors), outcomes and treatment data, genomic data (e.g., SNP, GWAS), and biospecimens (e.g., DNA, plasma, tumor tissue) and is stored securely within the ISC identified by person identification number only from the six participating sites. The website (<http://www.bcfamilyregistry.org/for-researchers/initiate-collaborations>) provides information about the data and biospecimens that are available and a detailed description of the BCFR, including information on the derived variables and data dictionaries. We have also contributed information on the data collected through our registry to the NCI Cancer Epidemiology Descriptive Cohort Database (<https://cedcd.nci.nih.gov/home.aspx>). Specific to genomic and genome-wide association study (GWAS) data, we comply with NIH Genomic Data Sharing (GDS) Policy, and will submit any large-scale human genomic data (e.g., GWAS, SNP arrays, epigenomic) to the National Center for Biotechnology Information's (NCBI) database of Genotypes and Phenotypes (dbGaP). Finally, we have published a manuscript on our open resource for collaborative research within the journal, Breast Cancer Research and Treatment, in 2009 to describe and promote use of the resource.

We have developed a system to promote and participate in external collaborations and consortium efforts. Resource sharing with external collaborators is done through the BCFR ISC which has implemented a tracking system for all data and biospecimen requests, including a web-based request and tracking system. This facilitates transparency and use by external investigators. All BCFR sites and the ISC follow established NIH guidelines for data sharing (<http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03032.html>). New requests for data, by submission of an application (please see following document), are made to the ISC which logs each request and generates the appropriate summary tables. Individual or aggregate level data, as requested and approved, are released only to investigators after an application has been approved and appropriate Data Use Agreements and IRB documents are in place. Biospecimens for approved projects are released by the BCFR biorepositories after appropriate Material Transfer Agreements are in place. Upon receipt of the data, we work closely with investigators on their preparation, analysis and presentation of the data to prevent any misuse and/or misinterpretation of the data. In addition, we have attended the NCI Cohort Consortium meeting to identify areas of synergism for research endeavors and promote collaborations.