

## Breast Cancer Cohort Study SOP for Data Request, Transfer and Storage

Includes Accepting Participants for the Approved Research Project “Breast Cancer Cohort Study.”

Please see “Greater Plains Collaborative Cooperative Medical Informatics Data Sharing and Network Infrastructure Agreement” for terminology

After all reviews and approvals specified in the above agreement, the following four data request work plans will ensue:

**Work Plan 1:** create the survey sample finder file. *This work plan is necessary to validate the cohort eligibility criteria are comparably defined and applied.*

- a. The Participant Honest Broker for each Accepting Participant will query its Research Repository as directed by the Breast Cancer Cohort Study: Include all people diagnosed with primary breast cancer (SEER Site Recode=Breast) diagnosed during the period of interest (5/1/2013 – 5/1/2014 at all sites except the University of Iowa and University of Nebraska where the start date is 1/1/2013) who are age 18+ at the time of diagnosis. The following variables needed for case selection will be returned in the query: encounter\_num (a system-generated number that designates the tumor to which all other tumor descriptors including date of diagnosis pertain), primary site (should be only breast), cancer identification (masked date of diagnosis, grade, diagnostic confirmation, morphology, sequence number, AJCC-7 stage, SEER Summary Stage 2000), and demographics (masked date of birth, sex, vital status, race, Spanish/Hispanic origin).
- b. Each Participant Honest Broker will freeze this “large cohort” by maintaining a link between the de-identified patient number *at the time of the query* and the actual medical record numbers.
- c. The Participant Honest Brokers will submit the requested information in the agreed-upon format to the GPC Honest Broker.
  - The information will include a deidentified individual level dataset containing necessary data elements required to confirm eligibility for the study and also quality check that the query had delivered expected results.
  - The GPC Honest Broker will exclude patients who are not eligible using the variables above: Sex not female, less than 18 years of age, prior cancer diagnosis (sequence number >1), first breast cancer not diagnosed during period of interest, breast cancer not microscopically confirmed, only tumor morphology is lobular carcinoma in situ, stage IV breast cancer, or known to be deceased.
- d. The GPC Honest Broker will store the Response in a REDCap database with attached files for retrieval on the GPC Hosted Services
- e. The Breast Cancer Cohort Study statisticians and epidemiologists will analyze the Response data for quality checking. Suspicious data patterns will result in communication, clarification, and data correction with the Accepting Participant
- f. They will then supply the de-identified patient number for these cases to Breast Cancer Cohort Study staff.
- g. From this list, the Breast Cancer Cohort Study statisticians and epidemiologists will create a “Survey Sample” finder file for each Accepting Participant.
  - This will be a random sample of sufficient size to generate the desired number of survey responses.
- h. The Participant Honest Broker will restore the link to the medical record/encounter number and verify that the cases in the Survey Sample meet study criteria.
  - The breast cancer survey site PI is responsible for verifying that each case meets study criteria. Most sites will accomplish this by arranging for their hospital’s tumor registrar to compare the list and diagnosis dates with breast cancer cases in the tumor registry.

- i. The Participant Honest Broker will retrieve name and mailing address and supply a file to the site breast cancer PI containing name, address, and de-identified patient number for mail merge. See survey mailing protocol for subsequent steps.

**Work Plan 2:** Data support for mailings *This work plan specifies how sites generate mailing list and conduct initial and non-respondent mailings and how response is tracked*

- a. The Breast Cancer Cohort Study will submit one Survey Sample finder file for each Accepting Participant to the GPC Honest Broker
- b. The GPC Honest Broker will submit the Survey Sample finder file to the appropriate Participant Honest Broker.
- c. The Participant Honest Broker will append a study ID number (a range of ID numbers will be specified by the Coordinating Center) and the MRN number to the Survey Sample finder file and retain this file for later data requests for consented patients.
- d. The Participant Honest Broker will match the Survey Sample finder file to name and address information to create a Mailing List file containing this information plus the study ID number.
- e. The Participant Honest Broker will submit the Mailing List file to the Participant Breast Cancer Cohort Study site PI or their designee for verification. It is strongly recommended that the sites work with their local tumor registrar to verify the list (i.e. that all are women diagnosed with breast cancer during the period of interest).
- f. Mailing
  - i. Sites will implement a mail merge, mailing and tracking protocol as directed by the Breast Cancer Cohort Study Coordinating Center.
  - ii. Participant Honest Broker or research assistant for site study PI will conduct a mail merge of the verified Mailing List file with the Coordinating Center-supplied materials (study booklet, site-specific cover letter, postage paid mailing envelope, incentive and reply envelope)
  - iii. Site research assistant stuffs envelopes and updates the site's tracking system (protocol includes double verification that name and address matches study ID on all enclosed materials, that incentive is enclosed, etc.)
  - iv. Coordinating Center reports study ID numbers of nonrespondents back to the sites
  - v. Sites make one re-mailing to nonrespondents
  - vi. If a participant included one or more pieces of information (i.e., date, printed, name, signature) on the medical record or future study consent documents - but not all 3 pieces of information, AND the participant did not provide their contact information in the booklet, the CC will contact the participating site and request that they contact the participant using the contact information in their mailing list for the associated participant study ID. The CC will provide materials for the sites to use in contacting the participant on a case by case basis. The same process will be implemented in the event that a booklet is received with 1 or more blank booklet pages when other pages were completed and there is no indication that the participation did not intend to complete the blank pages. This is because pages may have been stuck together and inadvertently left blank by the participant. Up to 3 call attempts will be made to contact the subject.

**Work Plan 3:** Store survey dataset *This work plan stores the survey data in the study database and uploads survey data for consented subjects back to the sites.*

- a. During data collection, the survey database will be stored on a secure database server at The UI College of Public Health.
- b. At the end of the study the Breast Cancer Cohort Study project manager will submit each Accepting Participant's survey dataset (for consenting subjects only) to the GPC Honest Broker for hosting on the GPC Hosted Services.

- i. Consenting subjects are those who have signed a consent form that indicates their answers will be housed in a repository for future uses that may not just be for breast cancer research. This is optional and some patients will choose to participate only in the survey.
- ii. The survey dataset will be a limited data set. It will not contain name, address, telephone, email address, or names and addresses of service providers. It will contain dates (date of birth, date of medical events, etc). It will also contain study ID number.
- c. If desired, the Participant Honest Broker for each Accepting Participant may download their consented subjects survey dataset. If they choose to do so:
  - i. The Participant Honest Broker will link the consented subjects' survey dataset to MRN number.
  - ii. The Participant Honest Broker will de-identify and transform the consented subjects survey dataset on their i2b2 staging server, and transfer the de-identified database to the de-identified server.

**Work Plan 4:** Request and obtain data for consented subjects. *This work plan is needed to complete the aims for the study and describes how the research team can access the complete study database for analysis.*

- a. The Breast Cancer Cohort Study Project Manager will provide one Consented Subjects finder file per Accepting Participant containing the study ID numbers for consented subjects to the Coordinating Center Honest Broker.
- b. The Coordinating Center Honest Broker will submit the Consented Subjects finder file to the GPC Honest Broker.
- c. The GPC Honest Broker will submit the finder file to the appropriate Participant Honest Broker
- d. The Participant Honest Broker for each Accepting Participant will query its Research Repository as directed by the Breast Cancer Cohort Study and submit the information in the agreed-upon format to the GPC Honest Broker. The information will include the study ID number.
- e. The GPC Honest Broker will create and store the Aggregated Response on the GPC Hosted Services.
- f. The dataset will be a limited data set. It will not contain name, address, telephone, email address, or names and addresses of service providers. It will contain dates of encounters and the study ID number.
- g. Only research team members will have access to the Aggregated Response dataset.