Chapter 6: Data and Biospecimen Sharing Policies and Procedures

6.1. DATA SHARING POLICY AND PLAN

6.1.1. Preamble

Consistent with the NIH mission to improve public health through research, the PCGC believes that the full value of genomic data to the public can be realized only if the genotype and phenotype datasets are made available as rapidly as possible to a wide range of scientific investigators. Rapid and broad data access is particularly important for genomic data because of the significant resources they require; the challenges of analyzing large datasets; and the extraordinary opportunities for making comparisons across multiple studies.

6.1.2. Policy

- The Centers of the PCGC will share genomic data produced through the PCGC award.
- Data sharing will be consistent with the NIH data-sharing policy for genome-wide association studies.
- The Centers of PCGC will abide by the PCGC Data Sharing Plan.
- Protection of research participants is a fundamental principle, underlying biomedical research. The PCGC and NIH is committed to responsible stewardship of data throughout the research process, which is essential to protecting the interests of study participants and to maintaining public trust in biomedical research.

6.1.3. Plan

The Centers of the PCGC will share genomic data produced through the PCGC award. The data will be shared through the NIH GWAS data repository (dbGaP). PCGC plans to release the data to qualified researchers who are approved to collaborate with the PCGC investigators. The availability of data for collaborators will be advertised on the PCGC web site. The data will be available through secure portal or on encrypted physical media.

The IRBs of the PCGC Centers have verified that:

- The submission of data to the NIH GWAS data repository and subsequent sharing for research purposes is consistent with the informed consent requested of the study participants from whom the data and specimens will be obtained
- The investigators' plan for de-identifying datasets is consistent with the standards outlined in the NIH GWAS policy
- It has considered the risks to individuals, their families, and groups or populations associated with data to be submitted to the NIH GWAS data repository and
- The genotype and phenotype data to be submitted will be collected in a manner consistent with 45 C.F.R. Part 46

For patients enrolled prospectively in PCGC-funded studies (e.g. CHD GENES) PCGC will share study documents and individual-level genotype and phenotype data. Study documents will include the study protocol, manual of operations, and case report forms. Genotyping data will include the genotypes as well as the intensity files used to call the genotypes. Phenotype data will include patient characteristics, diagnoses, and clinical outcome. Data will be shared for all of the consented study participants.

For patients enrolled retrospectively in PCGC studies (existing samples) PCGC will share summary level genotype and phenotype data consistent with historical consents that may or may not have addressed data sharing issues.

PCGC will share genotype and phenotype data once the genotyping data have been cleaned; cleaning is expected to be complete approximately 2 months after genotyping is finished. From the time the dataset is made available for access through the NIH GWAS data repository, a 12-month period of exclusivity will exist for PCGC investigators to submit analyses of genomic datasets for publication.

PCGC acknowledges the intellectual property and scientific publication elements of the NIH GWAS Policy.

6.1.4. Procedures

6.1.4.1. Data Storage

Internal to B2B: PCGC data will be stored in a central data repository and in the laboratories of PCGC investigators. The central data repository will be housed at one of the PCGC sites and will be overseen and managed by a bioinformatics group. Access to the central data repository will be limited to B2B investigators and Ancillary Study investigators (see Ancillary Study policy).

External to B2B: Analyzed and/or curated data will be stored in the Database of Genotypes and Phenotypes (dbGaP - http://www.ncbi.nlm.nih.gov/gap**),** housed at the National Center for Biotechnology Information, for the purposes of sharing with the wider scientific community and public**.**

6.1.4.2. Data Submission

Data management within PCGC and sharing with NIH data repository will protect rights and privacy of study participants. Specifics are spelled out in the IRB approved consent signed by participants. In cases where existing samples are used, sharing will be consistent with the consent under which those samples were obtained.

- Submissions to dbGaP should include the following:
- Study documents (i.e. manual of procedures, protocols, questionnaires, consent forms, etc.)
- Data dictionary (a description of measured variables with pointers to those parts of study documents that describe how variables were measured)
- Any other supporting documentation
- Phenotype, exposure, genotype, and pedigree data without identifiable information, created using a random, unique code whose key will be held by the submitting institution

Datasets will be made available to dbGaP as soon as appropriate quality control measures (as defined for each type of dataset) are complete. From the time the dataset is made available for access through dbGaP, a 12-month period of exclusivity will exist for PCGC investigators to submit analyses of genomic datasets for publication.

6.1.4.3. Data Access

Multiple tiers of data security will be employed based on the type (e.g. PHI, LDS, PII (Privacy Act) and level of risk associated with the data and any regulatory controls required.

Internal to B2B: Access to the PCGC central data repository will be limited to B2B investigators and ancillary study investigators (see Ancillary Study policy). Access to the data repository will be transparent to all B2B investigators and will be governed by the PCGC Steering Committee (SC). Prior

to accessing data for analysis, a brief proposal will be submitted to the PCGC SC for approval (see Publication Policy).

External to B2B: Qualified, external researchers (non-B2B, non-ancillary studies) may access the genotype and phenotype datasets submitted and stored at dbGaP. A Data Access Committee (DAC) will oversee access to PCGC data contained in dbGaP. Investigators who wish to access shared data must complete a Data Use Certification, which is reviewed and approved by the DAC. Membership of the DACs will include Federal staff with relevant expertise in areas such as the relevant particular scientific disciplines, research participant protection, and privacy.

dbGaP provides two levels of access: open and controlled, in order to allow broad release of non-sensitive data, while providing oversight and investigator accountability for sensitive data sets involving personal health information. Summaries of studies and the contents of measured variables as well as original study document text are generally available to the public, while access to individual-level data including phenotypic data tables and genotypes require varying levels of authorization.

6.1.5. Intellectual Property

It is the hope of the NIH that genotype-phenotype associations identified through NIH-supported and NIH-maintained genomic datasets and their obvious implications will remain available to all investigators, unencumbered by intellectual property claims. The NIH discourages premature claims on pre-competitive information that may impede research, though it encourages patenting of technology suitable for subsequent private investment that may lead to the development of products that address public needs. The NIH expects that NIH-supported genotype-phenotype data made available through the NIH GWAS data repository and all conclusions derived directly from them will remain freely available, without any licensing requirements, for uses such as, but not necessarily limited to, markers for developing assays and guides for identifying new potential targets for drugs, therapeutics, and diagnostics. The intent is to discourage the use of patents to prevent the use of or block access to any genotype-phenotype data developed with NIH support. The NIH encourages broad use of NIH-supported genotype-phenotype data that is consistent with a responsible approach to management of intellectual property derived from downstream discoveries, as outlined in the NIH's Best Practices for the Licensing of Genomic Inventions and its Research Tools Policy. The filing of patent applications and/or the enforcement of resultant patents in a manner that might restrict use of NIH-supported genotype-phenotype data could diminish the potential public benefit they could provide. Approved users and their institutions, through the execution of an NIH Data Use Certification, will acknowledge the goal of ensuring the greatest possible public benefit from NIH-supported GWAS.

6.1.6. Publications and Presentation

The maximum period of exclusivity for PCGC investigators is 12 months from the date that the genomic dataset is made available for access through dbGaP. Publication exclusivity is expected to extend to all forms of public disclosure, including meeting abstracts, oral presentations, and publicly accessible electronic submissions. The NIH expects all investigators who access genomic datasets to acknowledge the contributing investigators who conducted the original study, the funding organization(s) that supported the work, and the NIH data repository in all resulting oral or written presentations, disclosures, or publications of the analyses. Decisions about list of authors and co-authors should be made in accordance with generally accepted authorship guidelines.

6.2. BIOSPECIMEN SHARING POLICIES AND PROCEDURES

6.2.1. Preamble

Consistent with the NIH mission to improve public health through research, the Pediatric Cardiac Genomics Consortium (PCGC) believes that the full value of biospecimens can be realized only if made available to a wide range of scientific investigators. The PCGC encourages investigators in the field to use these materials and to foster collaborative research where appropriate. This Policy and Procedures applies only to the use of biospecimens in ancillary studies, i.e., studies other the endorsed PCGC studies, as defined in the PCGC Research Policy document. These ancillary studies can be proposed either by PCGC Investigators or by other investigators who agree to collaborate with a PCGC Investigator.

6.2.2. Policy

The Centers of the PCGC will share biospecimens collected through the PCGC award for use in ancillary studies.

Investigators requesting access to PCGC biospecimens will abide by the PCGC Ancillary Studies Policy and Procedures and the Publication and Presentation Policy and Procedures.

6.2.3. Procedures

6.2.3.1. Specimen Storage

PCGC biospecimens will be stored in a central Biorepository and in the laboratories of PCGC investigators. The central Biorepository will be overseen and managed by the PCGC SC or other designee.

6.2.3.2. Specimen Submission

Data management within PCGC and the Biorepository will protect rights and privacy of study participants. Specifics are spelled out in the IRB approved consent signed by participants. In cases where existing samples are used, sharing will be consistent with the consent under which those samples were obtained.

6.2.3.3. Specimen Access

Access to biospecimens for use in ancillary studies will be transparent to all PCGC investigators and will be governed by the PCGC SC or designee. Access to the PCGC biospecimens will be prioritized as follows:

- First priority is given to PCGC investigators of the five Centers.
- Second priority is given to non-PCGC investigators of approved ancillary studies (see Ancillary Study Policy and Procedures) or investigators of the four CvDC Centers, should there be an interest.

Investigators interested in accessing biospecimens should submit their proposal for an ancillary study to the PCGC SC for review (see Ancillary Study Policy and Procedures).

The following requirements for specimen access apply:

1. Collaboration Requirement: The PCGC would like to ensure that PCGC investigators are appropriately acknowledged. Collaboration with PCGC investigators (past or present) is highly encouraged and will be necessary as determined by the PCGC SC.

2. Patient Consent: The conditions under which the biospecimens will be shared will consider the privacy of the individuals from whom the biospecimens were obtained, what the informed consent permitted, and the merits of the proposed research.
3. IRB Approval: Investigators will be required to provide evidence of Institutional Review Board (IRB) approval of their proposed research plan prior to receipt of any biospecimens.
4. Funding: Investigators will be required to provide evidence of funding of the ancillary study prior to release of any biospecimens.
5. Material Transfer Agreement: Investigators granted access to PCGC biospecimens for use in an ancillary study must execute and adhere to the requirements of a Material Transfer Agreement.
6. Reporting of Findings and Sharing of Data: Investigators should provide a report of their findings within six months of completing the study. Any remaining biospecimens should be returned to the PCGC and all genetic analysis data should be provided to the PCGC Bioinformatics Core by that time.