**Model Data Sharing Plan for Human Genomic Data**

***Please modify this template as appropriate for your study, using our suggested items in bold and notes to guide your final plan. Do not submit the model without editing to describe your study and data and to remove our notes and guidance. Please contact the CTSI Research Facilitators (******ctsi@pitt.edu******) with any questions about using this template or preparing your Genomic Data Sharing Plan.***

**Data Types:**

This study involves the collection of (**Human, Non-Human, both Human and Non-Human)** data. Human data has been collected (**prospectively, retrospectively, both retrospectively and prospectively**). Data will be made (**publicly available, available through controlled access**). Genomic data to be made available includes (**sequence, transcriptomic, epigenomic, gene expression data**). Additionally, we will make available accompanying (**phenotypic, exposure**) data. Additional materials needed to interpret these data will also be shared, including (**study protocols, data collection instruments, data dictionary, pipeline metadata**). These materials will be available for unrestricted access.

**IRB Assurance of the Genomic Data Sharing Plan:**

Human data collected (**prospectively, retrospectively, both retrospectively and prospectively**) will be shared in accordance with the NIH genomics data sharing policy (NOT-OD-14-124). For prospectively collected data, we (**will, have**) obtained review of the University of Pittsburgh Institutional Review Board. Based on the consent used to collect these data, they may be shared (**broadly, with limitations**). Limitations to data sharing are provided in this link: <http://gds.nih.gov/pdf/standard_data_use_limitations.pdf>. For genomic data derived from (**retrospective clinical samples, human cell lines**), we will use data derived from (**samples, cell lines**) generated (**before, after, both before and after**) January 25, 2105. Informed consent for data derived from (**samples, cell lines**) generated before January 25, 2105 allows these to be shared (**broadly, with limitations**). We will obtain informed consent for future research use and broad data sharing for any data generated from (**retrospective samples, cell lines**) (**obtained, created**) after January 25, 2015.

Note: *For IRB approval, you will need a data security plan and resources to cover these expenses in the budget.*

**De-identification:**

In preparing human data for data-sharing, we will de-identify human genomic data and (**phenotypic, exposure**) data to ensure that the identities of research subjects cannot be readily ascertained. We will strip the data of identifiers according to the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. We will have random unique codes and will secure the key to the data with other study identifiers.

**Repository:**

**Include all of the following statements that apply:**

* We will make human data for controlled access available through (**name of repository**) – an NIH-designated repository.
* We will make human data for unrestricted access available through (**name of repository**).
* We will make non-human data available freely available through (**name of repository**).
* Level 2 sequence data (e.g. BAM files) will be made available through (**Sequence Read Archive**, **Cancer Genomics Hub**).

For human data, and regardless of where the study data are to be deposited, we will register the study in the database of Genotypes and Phenotypes (dbGAP), by the time that data cleaning and quality control measures begin, which is estimated to be approximately (**# of months**) after the start of funding.

**Data Submission and Release Timeline:**

For data submission, we expect to adhere to the following general schedule.

**Include all of the following statements that apply:**

**Non-Human**

* We will make non-human genomic data available no later than the date of publication.
* We will make non-human genomic data available before publication, in accordance with the requirements of our program.

**Human**

* We will make human genomic data available on a timeline similar to the schedule recommended in “Supplemental Information to the National Institutes of Health Genomic Data Sharing Policy”. Level 2 data including (**DNA** **sequence alignments, RNA expression profiling**) and Level 3 data including (**SNP, structural variant calls, expression peaks, epigenomic features**) will be made available following data cleaning and quality control at approximately (**# months**) after generation. Level 4 data indicating relationship of genomic data to phenotypic data or other biological states will be made available as they are completed.
* Phenotype or clinical data will be submitted to the NIH-designated data repository at the earliest opportunity, but no later than the date of Level 2 genomic data submission.
* For studies in which phenotype data collections are ongoing and/or may be regularly updated, data files will be submitted to NIH-designated data repositories as early as possible considering the practical needs for ensuring data accuracy. We expect that this corresponds to approximately (**# months**) after data generation.

**Data release**

We accept the proposed schedule for data release provided in the document entitled “Supplemental Information to the National Institutes of Health Genomic Data Sharing Policy”.

(See <http://gds.nih.gov/PDF/Supplemental_Info_GDS_Policy.pdf>)

**Appropriate Uses of the Data**

Based on the consent, human data may be used by others for general research use, biomedical use, and disease-specific use along with the following limitations:

(See <http://gds.nih.gov/pdf/standard_data_use_limitations.pdf>)

**Request for an Exemption to Data Submission**

Include a specific request for exemption if submission is not appropriate (e.g., consent for retrospective data does not allow sharing) and provide an alternative mechanism for use of the data.