

**www.nimhgenetics.org**

**Planning & Budgeting for**

**Data Submission to NRGR**

*To assist you in planning and budgeting for data submission to NRGR, we provide below a list of activities that occur at different phases of your award. The estimated number of hours may be higher or lower depending on the complexity of the data you will submit. The estimates assume proficiency with the clinical instruments being submitted and with handling tabular datasets (e.g. Excel).*

**Preplanning and Post-Award** (10 Hours; Y1 of award)

* Review and sign necessary documents verifying appropriate consents for data and sample sharing as directed by NIMH
* Prepare and submit Material Transfer Agreement with Rutgers University
* Modify consent form to include biorepository sharing language
* Define data sharing schedule
* Attend NRGR training on data submission procedures & confirm expected data to be submitted to NRGR
* Attend IBX training on biomaterial submission procedures
* Request NRGR accounts - <https://www.nimhgenetics.org/create-account>

**Develop Data Submission Workflow for NRGR Standard Submission Files** (15-20 Hours)

* **Review Data Dictionaries to be used for submission** <https://www.nimhgenetics.org/submit-your-data/submission-requirements>
  + For all submissions, we require the Submission (\_sub) file and the Diagnosis (\_dx) file at a minimum. For submissions with data at additional repositories (e.g. NDA, dbGaP), the Alternate ID (\_id) file is also required.
* **Develop procedure to extract data from local database into format specified by data dictionaries –** We suggest using the AutoQC submission tool to test & validate a subset of extracted data before the final year of your grant.
* **Prepare auxiliary material for data submission (e.g. public-facing study description)**

**Develop Data Submission Workflow for Item Level Assessments** (5-10 Hours per simple instrument, 15-30 Hours per complex instrument)

* **Prepare Data Dictionaries to be used for submission** <https://www.nimhgenetics.org/submit-your-data/submission-requirements>
  + For studies submitting item-level instrument data, we require the user to define a Phenotypic Data Dictionary (\_phen\_dd) for each assessment being submitted. In some cases, there may be an existing data dictionary available.
    - Simple Instrument (e.g. YBOCS): <100 fields, few different value types (e.g. all "integer, encoded"), no/few conditional requirements
    - Complex Instrument (e.g. SCID): >100 fields,  many different value types, more conditional requirements
* **Develop procedure to extract data from local database into format specified by data dictionaries –** We suggest using the AutoQC submission tool to test & validate a subset of extracted data before the final year of your grant.
* **Prepare auxiliary material for data submission (e.g. copy of assessment(s))**

**Submit Data** (5-20 Hours; Final Year)

* **Use AutoQC tool to Validate/Correct Errors and Submit** - Resolve any data quality issues identified by NRGR AutoQC Tool. Frequently users need to verify and correct biosample IDs and their corresponding Subject codes at the time of sample submission
* **Review and Resolve Post Submission QA/QC Issues** – After successful AutoQC submission, NRGR curators review the details of your submission and may request additional data fields, clarifications about diagnoses to assist in cross-study harmonization, or data on additional samples received and stored at the biorepository but absent from the initial submission.