Data Submission Office Hours Webinar

April 2023

National Institute of Mental Health

Purpose

- **Purpose:** Learn the general NDA Data Submission process during a Submission Cycle

- **Agenda:**
  - Who, what, when and how to submit data
  - Prerequisites to data submission
  - How to check if your Data Structures are ready for submission
  - Submission Templates
  - Mandatory Data Structures
  - Submission Exemptions

NIH National Institute of Mental Health

Who should be submitting data & when?

NDA expects you to submit data biannually by deadlines:
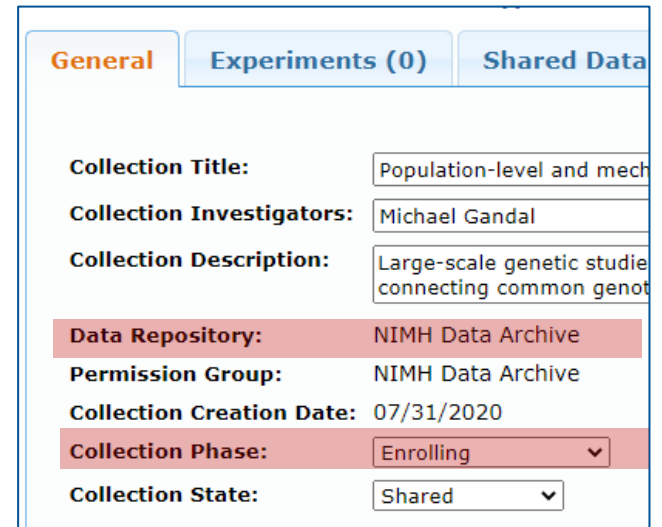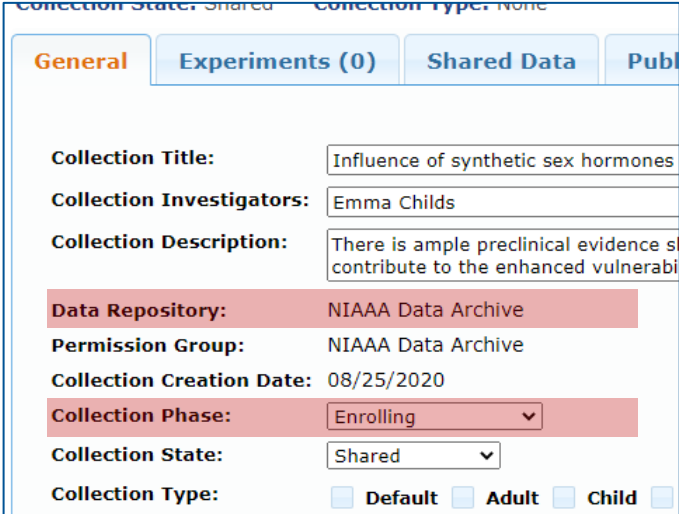
January 15 & July 15

1 Your NDA Collection is in the NIMH Data Archive data repository

2 Your Collection phase is in Pre-Enrollment, Enrolling or Funding Completed – Data Expected

3 The Data Expected list has an initial submission date set to the deadline date or earlier

General	Experiments (0)	Shared Data

Collection Title: Population-level and mech
Collection Investigators: Michael Gandal
Collection Description: Large-scale genetic studie connecting common genot
Data Repository: NIMH Data Archive
Permission Group: NIMH Data Archive
Collection Creation Date: 07/31/2020
Collection Phase: Enrolling
Collection State: Shared

∨ **Data Expected List: Mandatory Data Structures**

These data structures are mandatory for your NDA Collection. Please update the for the entire study.

Data Expected ⇕	Targeted Enrollment ⇕	Initial Submission ⇕
Research Subject and Pedigree ⓘ	46	10/01/2021

∨ **Data Expected List: Data Structures per Research Aims**

Who should be submitting data & when?

NDA expects you to submit data biannually by deadlines:

April 1 & October 1

1 Your NDA Collection is in the NIAAA Data Archive data repository

2 Your Collection phase is in Pre-Enrollment, Enrolling or Funding Completed – Data Expected

3 The Data Expected list has an initial submission date set to the deadline date or earlier

What do I need to start submitting data?

Data Submission Agreement (DSA)

Data Expected List

Permissions to Submit

Global Unique Identifiers (GUIDs)

NIH National Institute of Mental Health

Data Submission Agreement (DSA)

- Submit through the eDSA Tool
- Follow the instructions in Welcome to Data Sharing
- Access to your NDA Collection & Data Expected list

Data Expected List

- List of all Data Structures you'll submit data for
- Admin permissions required to edit

Permissions to Submit

- Submission permissions required to submit data
- Your PI manages permissions via the Permissions tab

Global Unique Identifiers (GUIDs)

- Required for all data submissions
- Access to NDA GUID Tool requested via the NDA Help Desk
- Submission permissions required

Global Unique Identifier (GUID)

Alphanumeric code used as an identifier for every research participant (ex. NDARRF67XH1)

What information do I need to create a GUID?

- First, middle and last name

- Sex at birth

- Date of birth

- City/Municipality of Birth

- *Recorded as it appears on the birth certificate

The NDA Data Sharing Terms and Conditions include the expectation that you will collect the data necessary to create real GUIDs.

NIH National Institute of Mental Health

pseudoGUIDs

pseudoGUIDs can be requested in place of GUIDs by emailing the NDA Help Desk:

- Detailed justification for the use of pseudoGUIDs
- # of pseudoGUIDs needed
- Whether pseudoGUIDs will be needed for more than one submission cycle

What if I obtain the information needed to create a GUID, but I've already used a pseudoGUID?

Promote your pseudoGUID using the NDA GUID Tool linking the pseudoGUID to a real GUID

NIH National Institute of Mental Health

What do I need to start submitting data?

Data Submission Agreement (DSA)

- Submit through the eDSA Tool
- Follow the instructions in Welcome to Data Sharing
- Access to your NDA Collection & Data Expected list

Data Expected List

- List of all Data Structures you'll submit data for
- [NDA Data Dictionary](#)
- Admin permissions required to edit

Permissions to Submit

- Submission permissions required to submit data
- Your PI manages permissions via the Permissions tab

Global Unique Identifiers (GUIDs)

- Required for all data submissions
- Access to NDA GUID Tool requested via the NDA Help Desk
- Submission permissions required

17

NIH National Institute of Mental Health

Data Submission

What data should be submitted?

- Please submit all data you have collected.

- You **DO NOT** need to submit data for all the Data Structures.

Should data be submitted cumulatively?

- Clinical and phenotypic data → *cumulative* submissions
  - Your first submission includes 5 subjects at Timepoint A
  - You collect 10 new subjects at Timepoint A & original 5 subjects at Timepoint B
  - Include the 5 subjects Timepoint A **&** B + 10 new subjects for Timepoint A

- Imaging, EEG, Omics data → non-cumulative submissions

NIH National Institute of Mental Health
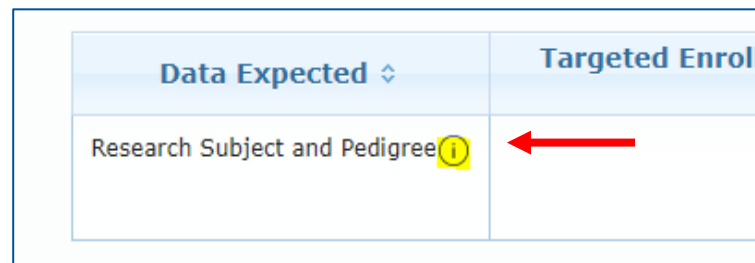
Data Submission cont.

Should I submit any data if NDA is still working on your Data Structures?

- ## Yes!

- You should submit data for Data Structures that are approved and are ready for submission!

How do I know if my Data Structure is ready?

Check the information ("i") icon next to the data structure names via your Data Expected tab!
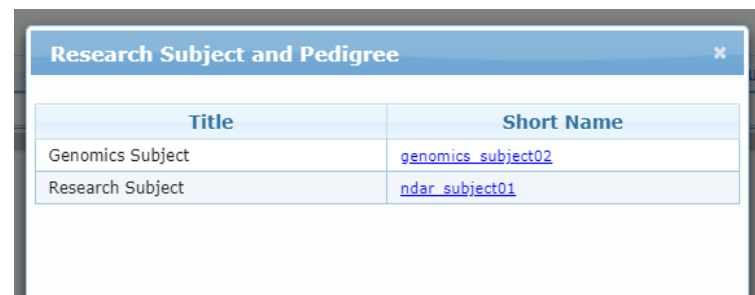
Structure not yet defined: Data curation is still preparing your Data Structure

If the pop-up provides the title and short name of your data structure: Your data structure is ready for submission!

NIH National Institute of Mental Health

Download Your Submission Template

Submission Template

- Do **not** change the first two rows.

- You can *hide* columns. Do **not** delete any columns.

- Each row = 1 subject at 1 timepoint
  - Example: 1 subject at 3 timepoints = 3 rows

National Institute of Mental Health

Submission Template

Required Data Elements

- Required data elements must have a value.

- Conditional data elements should have a value if applicable.

- Recommended and optional data elements can be left blank if not applicable.

ElementName	DataType	Size	Required
subjectkey	GUID		Required
src_subject_id	String	20	Required
interview_date	Date		Required
interview_age	Integer		Required
sex	String	20	Required

ElementName	Size	Required	Condition	Eleme
zygosity	255	Conditional	twins_study == 'Yes'	Zygos

ElementName	Size	Required	C
multibirthid		Recommended	
tourettes_disorder		Recommended	
sample		Recommended	
race_other	100	Recommended	
respond		Recommended	
date_confirmed_phenotype		Recommended	
pos_years		Recommended	
position1	150	Recommended	

24

National Institute of Mental Health

Required Data Elements

- If you did not collect a required data element, review ValueRange & Notes for N/A or null values.

ElementName	Size	Required	ValueRange	Notes	A
sex	20	Required	M;F; O; NR	M = Male; F = Female; O=Other; NR = Not reported	c
zygosity	255	Conditional	monozygous; dizygous; trizygotic; 999	999=unknown	m
reg_edu		Recommended	1::34; 999	1=None; 2=Preschool; 3=Kindergarten; 4=1st grade; 5=2nd grade; 6=3rd grade; 7=4th grade; 8=5th grade; 9=6th grade; 10=7th grade; 11=8th grade; 12=9th grade; 13=10th grade; 14=11th grade; 16=High School Graduate; 17=GED; 18=Some college credit but less than 1 year; 19= Some college, no degree; 20=Associates degree (e.g., AA, AS); 21=Bachelor's degree (e.g., BA, AB, BS); 22=Masters degree (e.g., MA, MS, MEng, MEd, MSW, MBA); 23=Professional degree (e.g., MD, DDS, DVM, LLB, JD); 24=Doctorate degree (e.g., PhD, EdD); 25=One or more years of college at a 2-year program, no degree; 26=One year of college at a 4-year program, no degree; 27=Two years of college at a 4-year program, no degree; 28=Three years or more of college at a 4-year program, no degree; 29= Trade or vocational school; 30=Eighth grade or less; 999=Unknown	
brain_injury		Recommended	1::3	1=Yes; 2=No; 3=NA	

25

5 Required Data Elements in All Structures

ELEMENT NAME	DATA TYPE	SIZE	REQUIRED	DESCRIPTION

How do I submit data?

Via the [Validation and Upload Tool](#)!

- Submission permissions required

- HTML Validation and Upload Tool
  - Cumulative data (e.g. clinical and phenotypic data)
  - [Tutorial for the HTML Validation and Upload Tool](#)

- Python Client Validation and Upload Tool
  - Non-cumulative imaging files
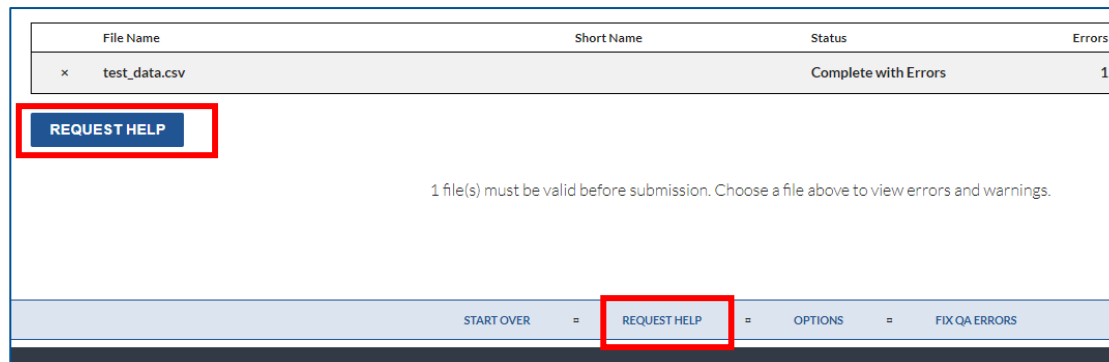 - Files over 2.5 GB

*You **must** use the Submission Templates from the NDA Data Dictionary and/or be provided by a Data Curator to submit your data.

NIH National Institute of Mental Health

Validation and Upload Tool

Validates data against the NDA Data Dictionary	Ensures data is harmonized correctly

- Use it as your personal NDA data manager
- Warnings vs Errors
  - Errors **must** be resolved
  - Users **can** submit with warnings present
- Need help? Click "Request Help"

What if none of your Structures are ready?

- If you are waiting on *NDA* to complete your data structures where you have ready-to-submit data, **don't worry!**

- There will always be at least one ready-to-submit data structure.

National Institute
of Mental Health

Mandatory Data Structures

Data Expected List: Mandatory Data Structures

These data structures are mandatory for your NDA Collection.

Data Expected ⇕
Research Subject and Pedigree ⓘ
NIMH CDE - GAD-7 ⓘ
NIMH CDE - PHQ-9 ⓘ
NIMH CDE - DSM-5 Crosscutting Adults ⓘ
NIMH CDE - WHODAS 2.0 ⓘ

Research Subject & Pedigree (RS&P)

- ndar_subject01
- genomics_subject02 (omics only)

NIH Common Data Elements (CDE) (NOT-MH-20-067)

National Institute of Mental Health

Mandatory Data Structures

- **Update** the targeted enrollment number.

- **Submit** data cumulatively.

Don't have any data to submit / Structures not ready?

Submit a submission exemption via the Data Expected tab of your NDA Collection!

Types of Exemptions

A. <u>No New Data This Period</u>: No new data to submit

B. <u>Deadline Extension</u>: Data submission expected after the submission deadline

C. <u>No New Data - Change Phase to Data Analysis</u>: Data submission is completed for the entire grant

Received QA Results?

If QA errors were identified from a previous submission cycle, do I still submit data?

- Please [correct the QA errors](correct%20the%20QA%20errors) and resubmit your data before submitting any new data.

- Follow the instructions in your QA results notification.

NIH National Institute of Mental Health

More Questions?

If you have further questions or would like a copy of the slide deck, please email

the NDA Help Desk at NDAHelp@mail.nih.gov.

All information from the presentation and more can be found on our Submission Cycle FAQ: https://nda.nih.gov/nda/submission-cycle-faq.html.

Thank you for joining!

National Institute of Mental Health