



NIMH Data Archive

How to Submit Corrected Data to Fix QA Errors

Last Updated: October 2022

Introduction

After data submission cycles, the NIMH Data Archive (NDA) performs [Quality Assurance and Quality Control \(QA/QC\) checks](#) across records, ensuring basic information is accurate. If any errors are found in submissions, NDA notifies the Principal Investigator(s) (PI), via email, with a QA errors report containing a summary and detailed QA/QC issues. PIs should correct each error using [Detailed QA/QC Issue Resolutions](#).

The following step-by-step instructions provide guidance on how to submit corrected data to fix identified QA errors.

Table of Contents

Introduction	2
Table of Figures	3
Resubmit using the HTML Validation and Upload Tool	4
Resubmit using the Python NDA-Tools Command Line	14

Table of Figures

Figure 1: NIMH Data Archive Homepage.....	4
Figure 2: Tools Menu - Validation & Upload Tool Menu Item	4
Figure 3: HTML Validation and Upload Tool Link	5
Figure 4: HTML Validation and Upload Tool Landing Page.....	5
Figure 5: NDA RAS Sign-In Page	6
Figure 6: Submissions with QA Errors Table	6
Figure 7: Select Submission & Click Submit Button.....	7
Figure 8: Filter Submission ID Field	7
Figure 9: List of All Data Structures in Selected Submission	8
Figure 10: Choose Files Link.....	8
Figure 11: Choosing File Dialog	9
Figure 12: Revalidate Link.....	9
Figure 13: Number of Rows in Data Structure and Uploaded Data File	10
Figure 14: Next: Build Package Button	10
Figure 15: Build Package Step	11
Figure 16: Edit Button	11
Figure 17: Selecting Acknowledgement Check Box and Submit Data Button	12
Figure 18: Submission Processing - Submission in Progress Message	12
Figure 19: Resubmission Complete Confirmation Message	13
Figure 20: Opened Terminal and Location of User's settings.cfg File	14
Figure 21: Example of a User's settings.cfg File, showcasing Username and Password Locations.....	14
Figure 22: NDA Profile Username Field and Reset Password Button.....	15
Figure 23: Example of Submission ID returned by the Python tool	15
Figure 24: Submission ID under Submission Tab of NDA Collection	15
Figure 25: Submission in Column C of QA Results Report	15
Figure 26: Example of Files Validating for Resubmission	16
Figure 27: Summary of Resubmission & Prompt to Continue	17
Figure 28: Successful Resubmission Confirmation Message	17
Figure 29: Unauthorized Request Message	18
Figure 30: Resubmission Error - Submission Already Replaced	18

Resubmit using the HTML Validation and Upload Tool

1. Navigate to the NIMH Data Archive (NDA) homepage: <https://nda.nih.gov/>

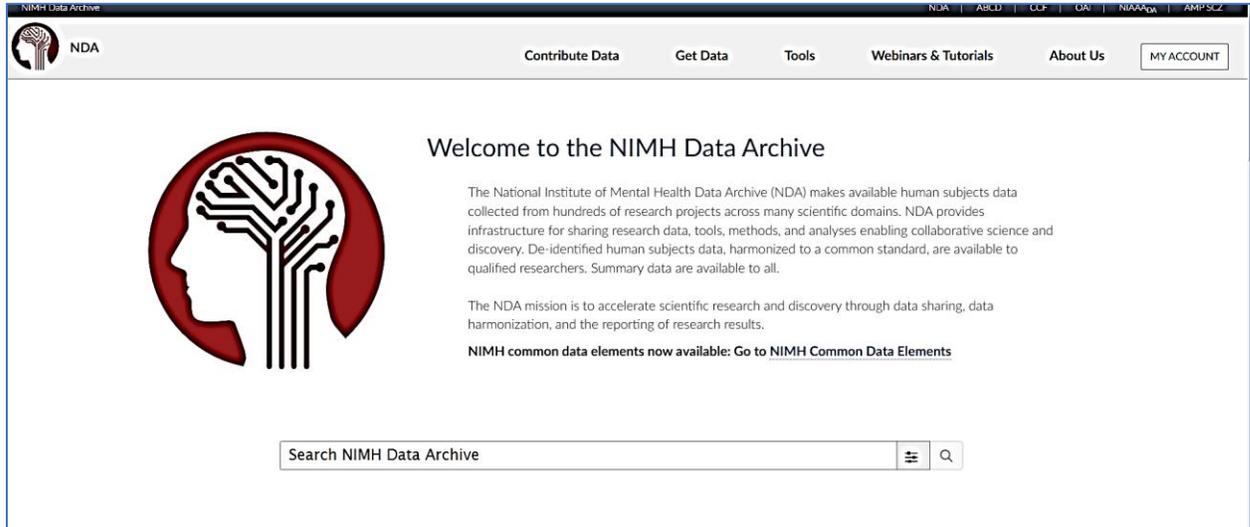


Figure 1: NIMH Data Archive Homepage

2. Navigate to "Tools" > "Validation & Upload Tool"

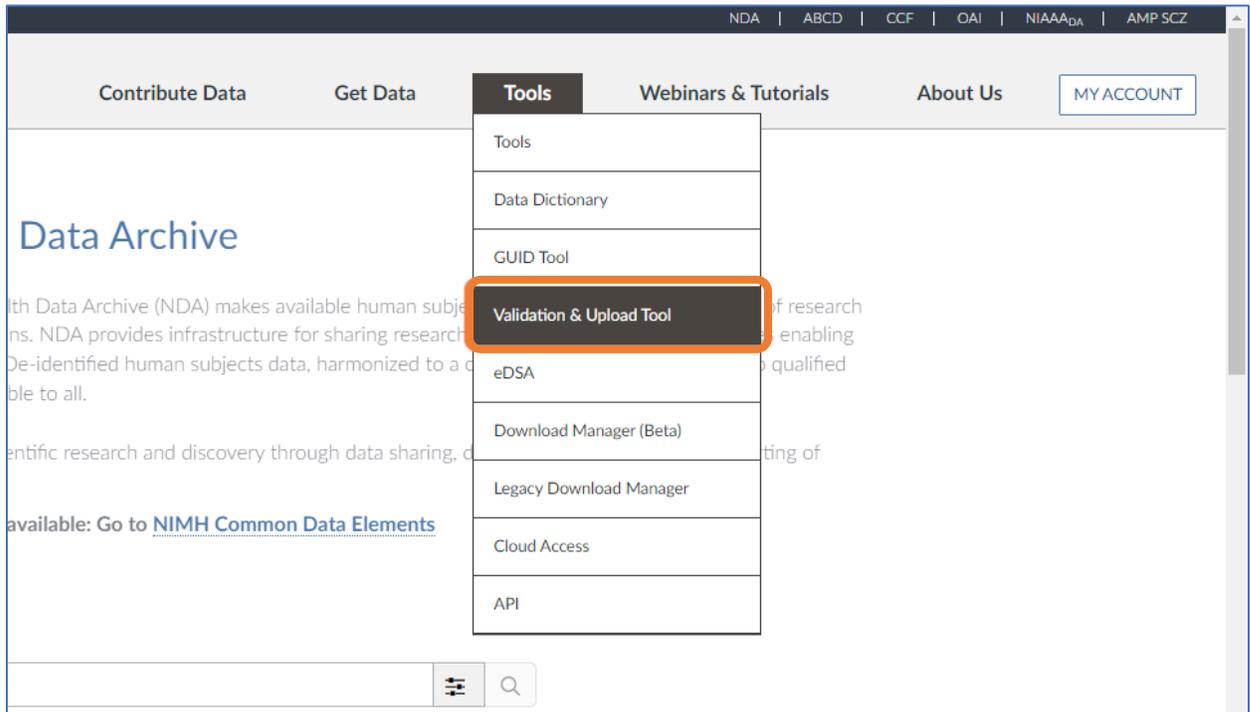


Figure 2: Tools Menu - Validation & Upload Tool Menu Item

3. Select "Launch the HTML Validation and Upload Tool"

AWS-RDS database. There are currently three different versions of the tool. Please review them below to determine the appropriate one for your situation. If in doubt, please use the first option: the HTML version.

HTML Validation and Upload Tool

This version of the tool allows you to use it a webpage, validating the quality of your data and upload it directly through your web browser. Chrome, Firefox, Safari, and recent versions of Internet Explorer are supported

Launch the HTML Validation and Upload Tool

Python Validation and Upload Tool

The tool is also available as a [Python package](#).

Figure 3: HTML Validation and Upload Tool Link

4. On the HTML Validation and Upload Tool landing page, click the "Fix QA Errors" button.

NIMH Data Archive

NDA ABCD CCF OAI NIAAData

NDA Data Validation and Upload Tool

Validate Data Associate Files Build Package Submit Data

Welcome to the Validation Tool HTML Client

Welcome to the NIMH Data Archive (NDA) Validation and Upload Tool. You can use this tool to ensure your data files are harmonized to the [NDA Data Dictionary](#), then upload them to your Collection, or resume an in-progress or interrupted validation/submission. To complete a successful submission you'll need to have your CSV submission files and associated source files (e.g. Imaging, EEG, Omic files) in a working directory, and an NDA account associated with your Collection(s).

Click below to begin, and [contact us](#) with any questions.

If you have a **submission in progress** please click **Resume Incomplete Submission** button to complete the submission; otherwise please click **Create New Submission** button.

RESUME INCOMPLETE SUBMISSION CREATE NEW SUBMISSION

START OVER REQUEST HELP OPTIONS FIX QA ERRORS

Figure 4: HTML Validation and Upload Tool Landing Page

- If you are not already signed into NDA, sign into your NDA account using your [Research Auth Service \(RAS\) Identity](#) (NIH, eRA Commons, or Login.gov).

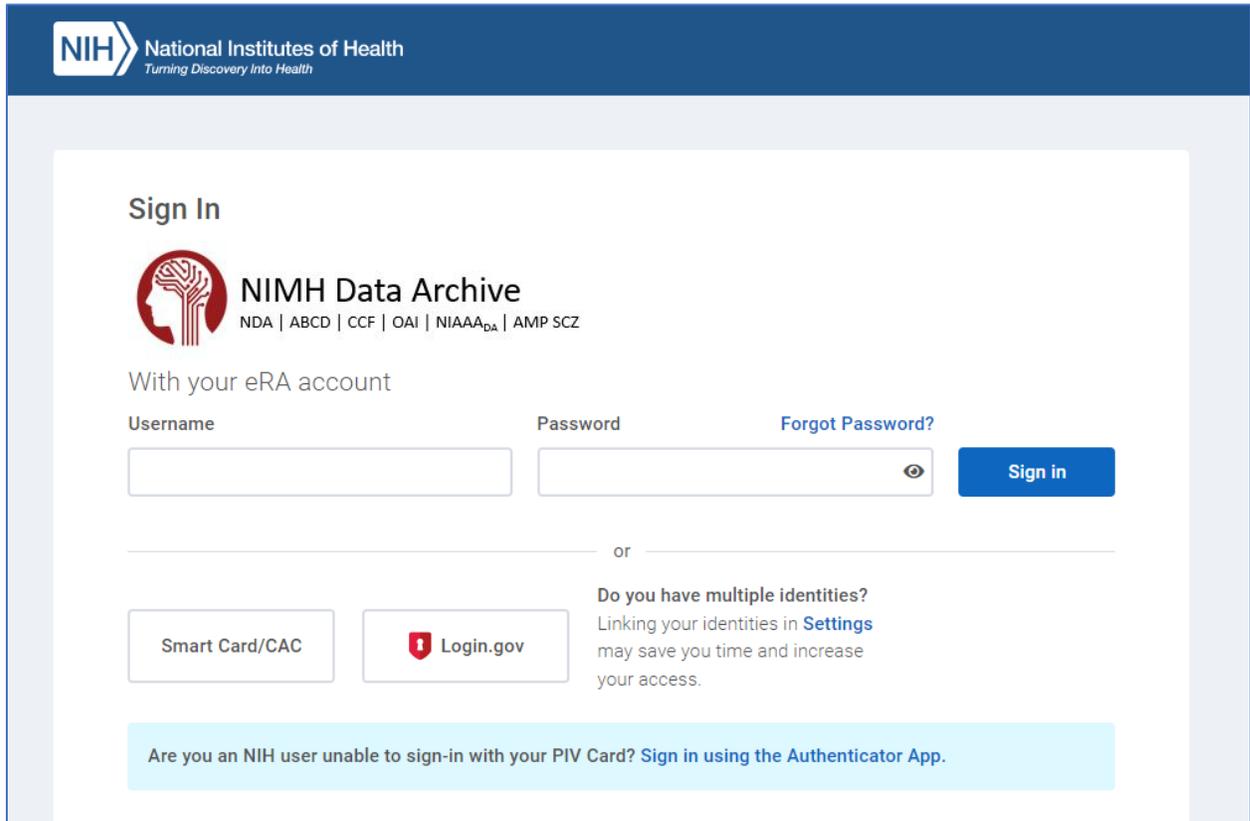


Figure 5: NDA RAS Sign-In Page

- Once logged in, the table displays all the QA errors found in previous data submissions for all NDA Collections you have access to.

NDA Data Validation and Upload Tool

Filter _____

Please select the submission for which you would like to fix QA errors

Select	Collection ID	Submission ID	Dataset Title	Dataset Description	Submission Status	Dataset Created Date	Dataset Modified Date
<input type="checkbox"/>	3151	23850	HBCC dataset-1	Test Submission 1	Upload Completed	Oct 24, 2019, 12:13:00 AM	
<input type="checkbox"/>	2373	31644	U54 Human Subjects Data	Test Submission 2	Upload Completed	Apr 20, 2020, 12:22:31 PM	
<input type="checkbox"/>	2270	35464	fMRI data	Test Submission 3	Upload Completed	Jan 19, 2021, 11:37:56 AM	

Figure 6: Submissions with QA Errors Table

7. Select one submission you have corrected data for and click "Submit."

Filter
45700

Please select the submission for which you would like to fix QA errors

Select	Collection ID	Submission ID	Dataset Title	Dataset Description	Submission Status	Dataset Created Date
<input checked="" type="checkbox"/>	17	45700	Resubmission Regression	Test Submission	Upload Completed	Dec 10, 2021, 11:12

SUBMIT

Figure 7: Select Submission & Click Submit Button

- a. You can also filter the table listings using the Submission ID in your QA errors report.

Validate Data

Associate Files

Filter
45700

Please select the submission for which you would like to fix QA errors

Select	Collection ID	Submission ID	Dataset Title	Dataset Description
--------	---------------	---------------	---------------	---------------------

Figure 8: Filter Submission ID Field

8. It's time to revalidate your corrected data files.
 - a. Only resubmit data for the **data structures that have identified QA errors** in the selected submission.
 - i. Example: Your selected submission contains 3 different data structures. Only 1 of the 3 data structures were identified to contain QA errors in your QA errors report. You should only resubmit data for the 1 data structure that contained QA errors.
 - ii. Refer to your QA errors report to confirm the data structures that have QA errors and need resubmitting.
 - iii. **New data should never be submitted until all QA Errors have been resolved.**
 - iv. Please note: The tool provides all data structures contained in the selected submission.

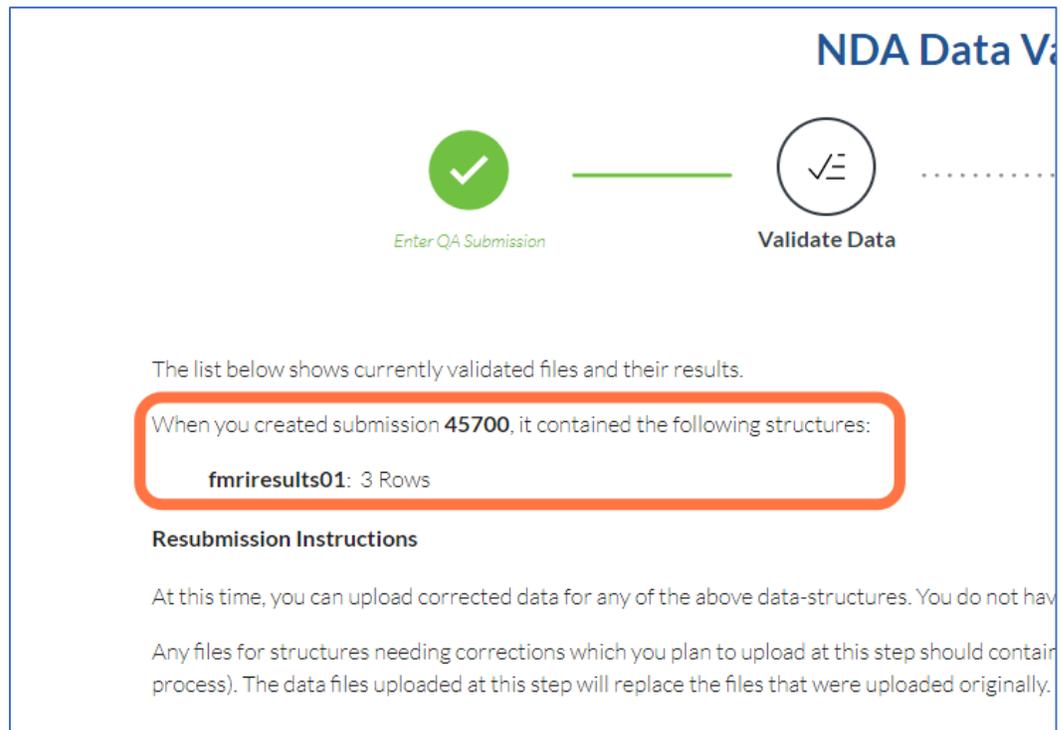


Figure 9: List of All Data Structures in Selected Submission

- b. Select “Choose Files” and choose your corrected data file(s). Alternatively, you can drag and drop your files.

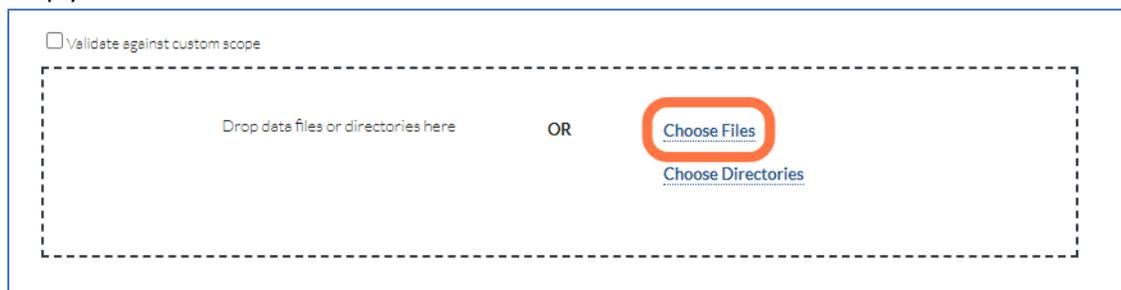


Figure 10: Choose Files Link

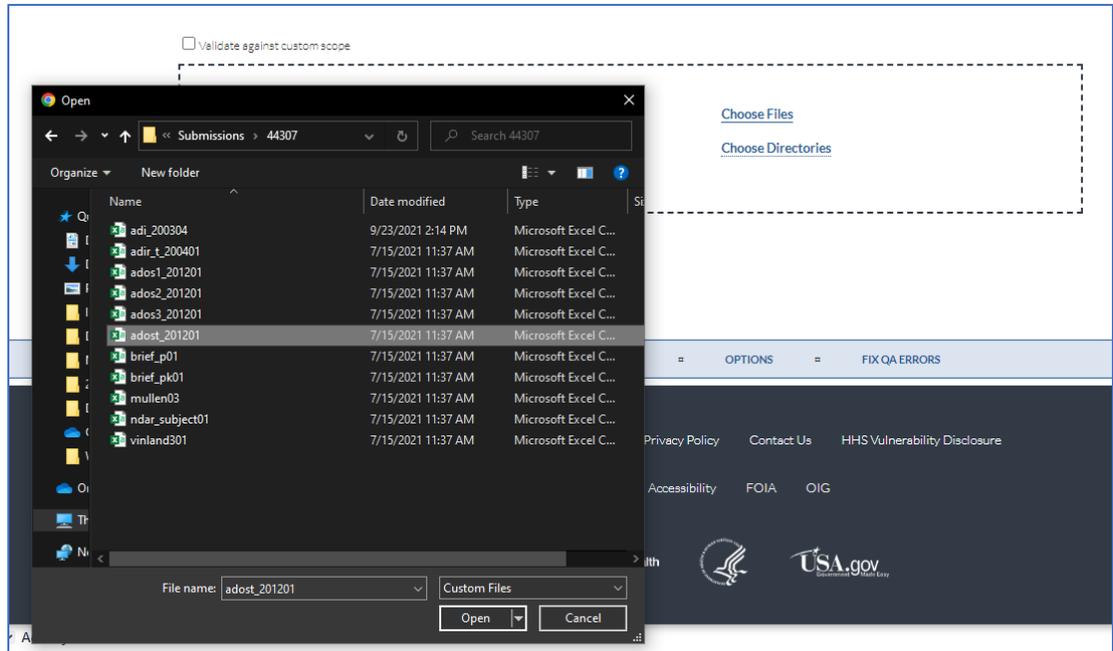


Figure 11: Choosing File Dialog

9. When you have selected your file(s) for resubmission,
 - a. Make sure there are zero errors. If any errors are found, you should correct and save your data file locally. Then click “Revalidate” to implement your corrections in the tool.
 - i. Alternatively, you can remove the file from the tool. Correct your data, then reupload your corrected data files.

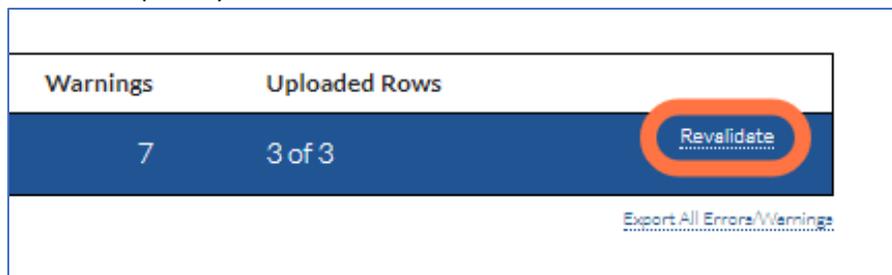


Figure 12: Revalidate Link

- b. You can continue with resubmission even if warnings are found.

- c. Your corrected data file **must have the exact number of rows** as indicated in the original submission.

When you created submission **45700**, it contained the following structures:

fmirresults01: 3 Rows

Resubmission Instructions

At this time, you can upload corrected data for any of the above data-structures. You do not have to upload files for a data-structure if no data needed to be corrected for it.

Any files for structures needing corrections which you plan to upload at this step should contain **ALL** of the rows from the original submission (minus duplicates which were detected during process). The data files uploaded at this step will replace the files that were uploaded originally.

Do **NOT** just submit the rows that needed to be corrected, as this will result in missing data from your collection.

Validate against custom scope

Drop data files or directories here

OR

[Choose Files](#)
[Choose Directories](#)

Selected file status
fmirresults01-small.csv: Potential data file extension detected.

File Name	Short Name	Status	Errors	Warnings	Uploaded Rows
x fmirresults01-small.csv	fmirresults01	Complete with Warnings	0	7	3 of 3

Figure 13: Number of Rows in Data Structure and Uploaded Data File

- d. Once you have confirmed this, click “Next: Build Package”

Errors	Warnings	Uploaded Rows	
0	7	3 of 3	Revalidate

[Export All Errors/Warnings](#)

NEXT: BUILD PACKAGE

Figure 14: Next: Build Package Button

10. The information entered in the original submission will display on the Build Package step.
 - a. Check to ensure the information shown is for the submission you intend to correct.
 - b. Once confirmed, click “Build Package.”

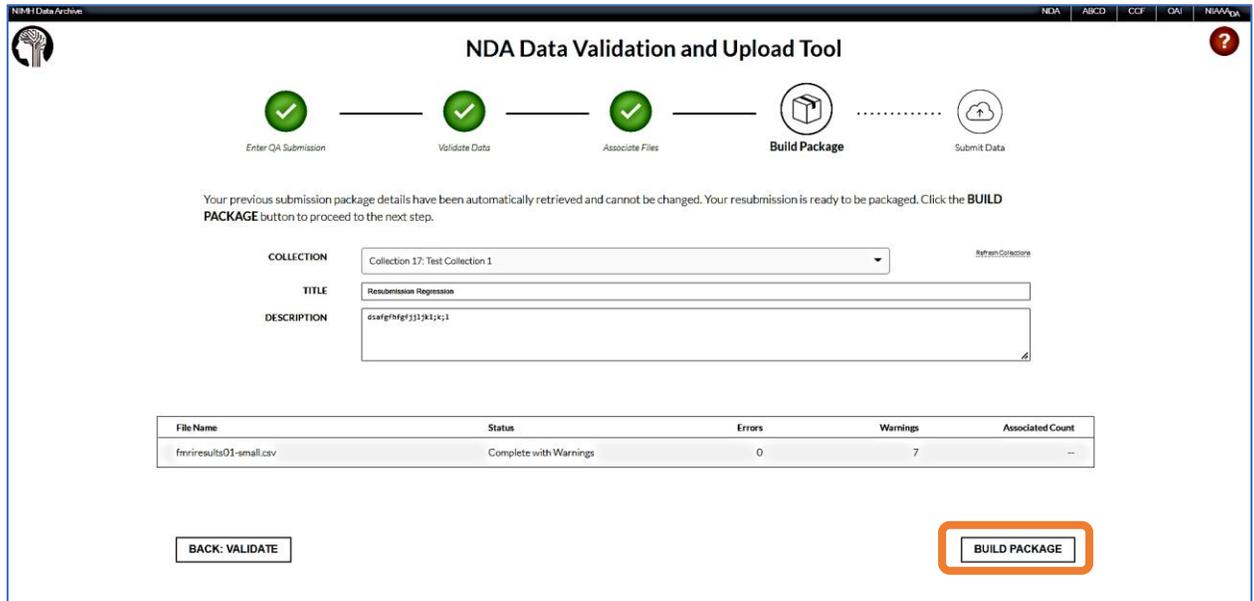


Figure 15: Build Package Step

11. The last step in the tool is the Submit Data step. Verify all the information in the Resubmission Summary is correct.
 - a. If any information needs to be changed, click “Edit”.

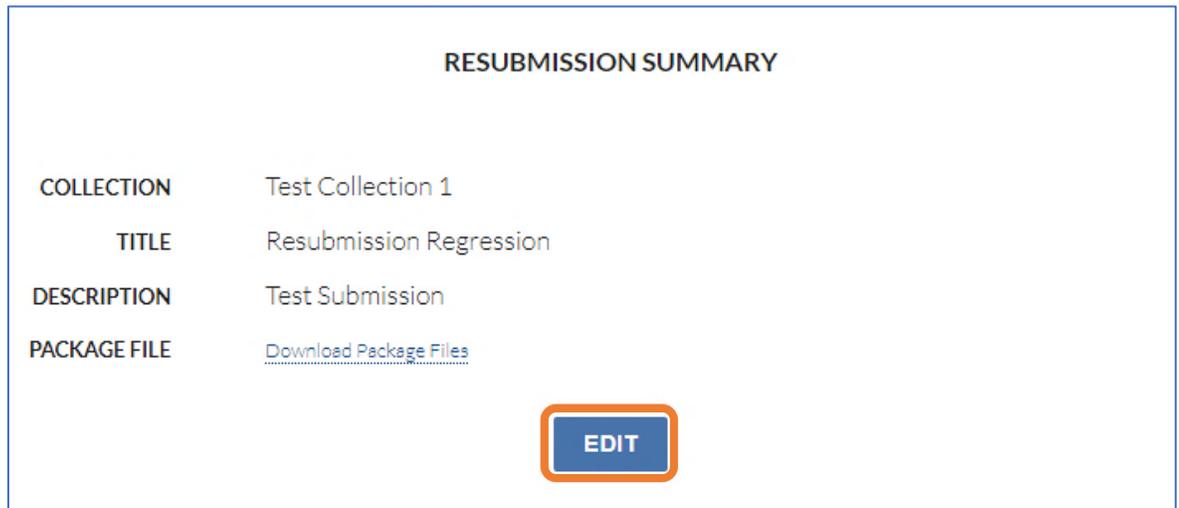


Figure 16: Edit Button

- b. Once the information is correct, confirm your submission does not include any Personally Identifiable Information (PII) by checking the acknowledgment check box.
- c. Click "Submit Data" to submit your corrected data files.

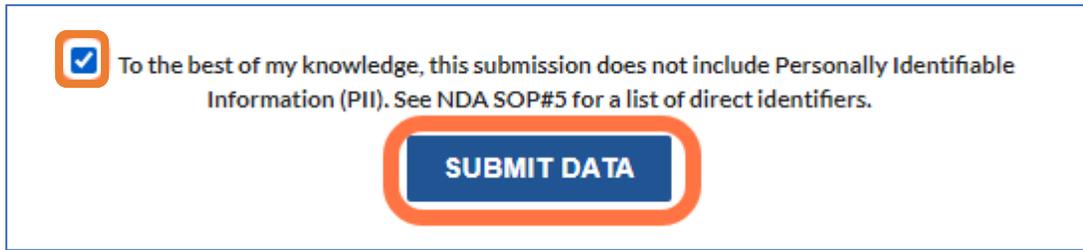


Figure 17: Selecting Acknowledgement Check Box and Submit Data Button

12. At this point, your submission will process.

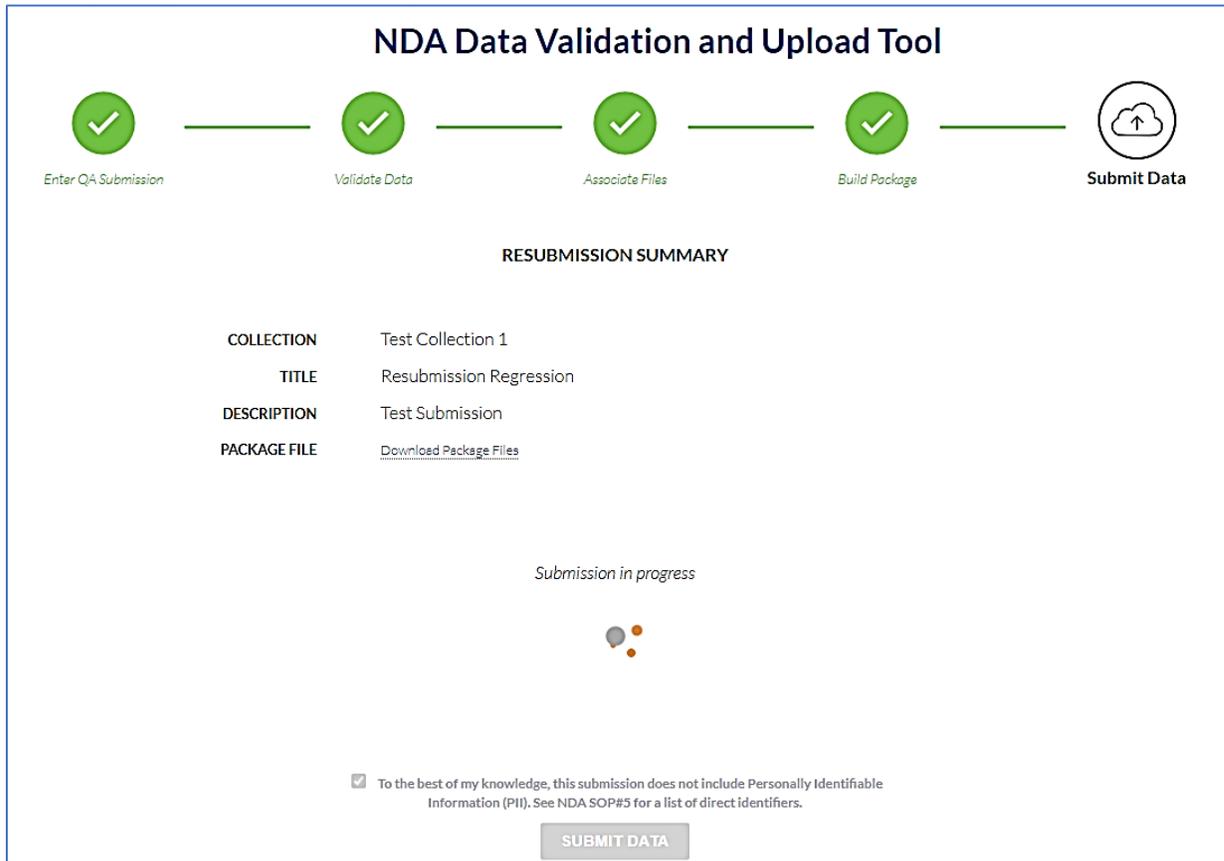


Figure 18: Submission Processing - Submission in Progress Message

13. Once your submission is fully processed, you'll receive a confirmation message informing you your submission was successfully submitted.
- Click Start Over to visit the [HTML Validation and Upload Tool](#) landing page where you can create a new submission, resume an incomplete submission, or resubmit data to correct QA errors again.

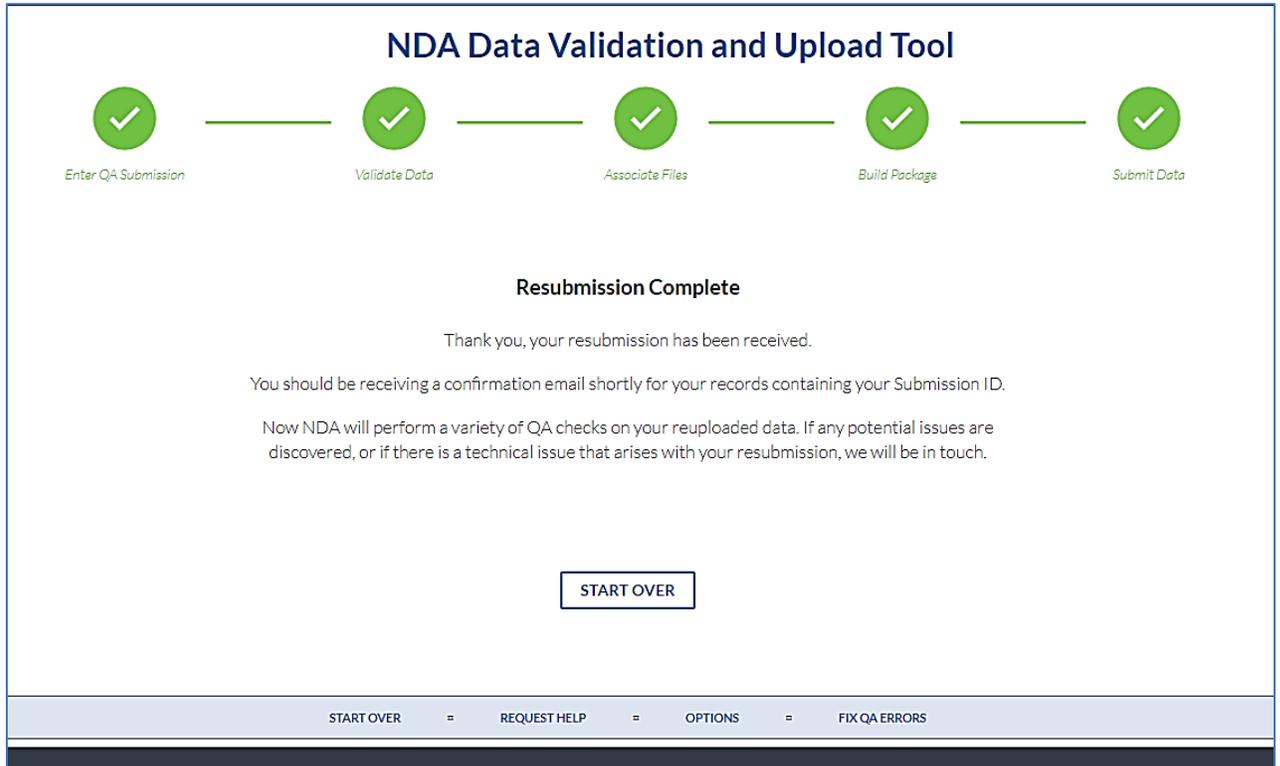
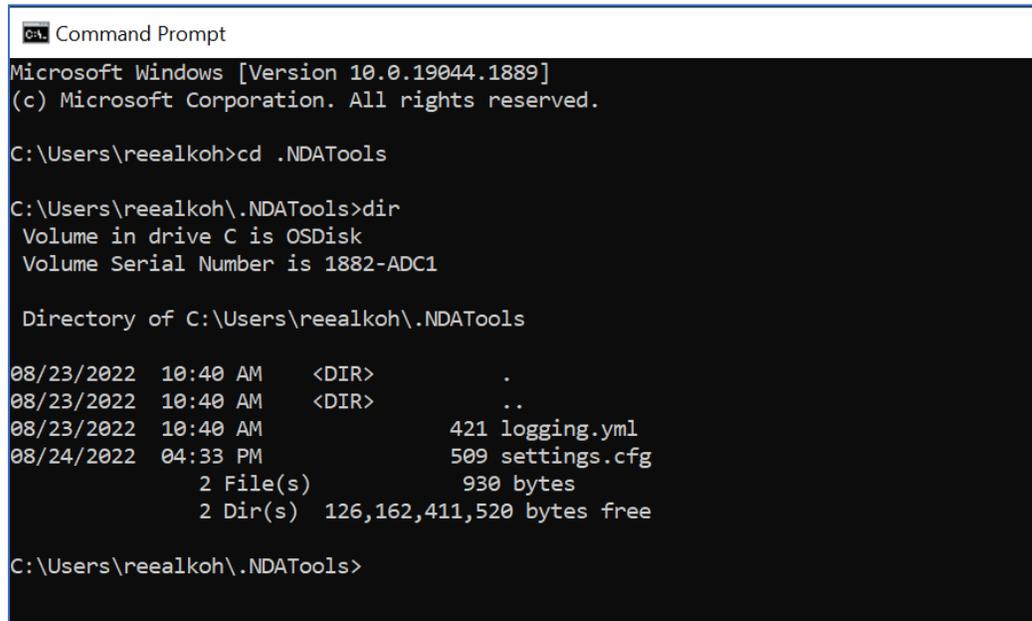


Figure 19: Resubmission Complete Confirmation Message

Resubmit using the Python NDA-Tools Command Line

NOTE: Only resubmit corrected data using this resubmission method if the data was *originally* submitted through the [Python NDA-Tools Command Line](#).

1. Open the terminal and directory previously used to upload their submission `NDATools\settings.cfg` file.



```
Command Prompt
Microsoft Windows [Version 10.0.19044.1889]
(c) Microsoft Corporation. All rights reserved.

C:\Users\reealkoh>cd .NDATools

C:\Users\reealkoh\.NDATools>dir
Volume in drive C is OSDisk
Volume Serial Number is 1882-ADC1

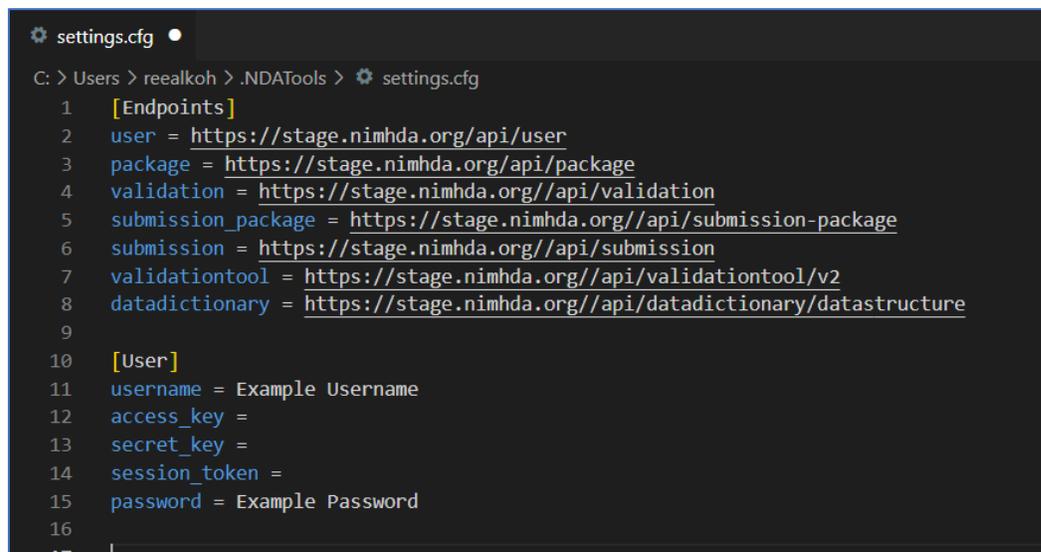
Directory of C:\Users\reealkoh\.NDATools

08/23/2022  10:40 AM    <DIR>          .
08/23/2022  10:40 AM    <DIR>          ..
08/23/2022  10:40 AM                421 logging.yml
08/24/2022   04:33 PM                509 settings.cfg
           2 File(s)                930 bytes
           2 Dir(s) 126,162,411,520 bytes free

C:\Users\reealkoh\.NDATools>
```

Figure 20: Opened Terminal and Location of User's settings.cfg File

2. Review, edit, and confirm your NDA username and password in the `NDATools\settings.cfg` to confirm you are connected to your NDA account.



```
settings.cfg
C: > Users > reealkoh > .NDATools > settings.cfg
1  [Endpoints]
2  user = https://stage.nimhda.org/api/user
3  package = https://stage.nimhda.org/api/package
4  validation = https://stage.nimhda.org//api/validation
5  submission_package = https://stage.nimhda.org//api/submission-package
6  submission = https://stage.nimhda.org//api/submission
7  validationtool = https://stage.nimhda.org//api/validationtool/v2
8  datadictionary = https://stage.nimhda.org//api/datadictionary/datastructure
9
10 [User]
11 username = Example Username
12 access_key =
13 secret_key =
14 session_token =
15 password = Example Password
16
17
```

Figure 21: Example of a User's settings.cfg File, showcasing Username and Password Locations

- a. You **must** use the username found in your NDA profile and your NDA Tools password.

To set or reset your NDA Tools password, click "Reset Password" in your NDA profile.

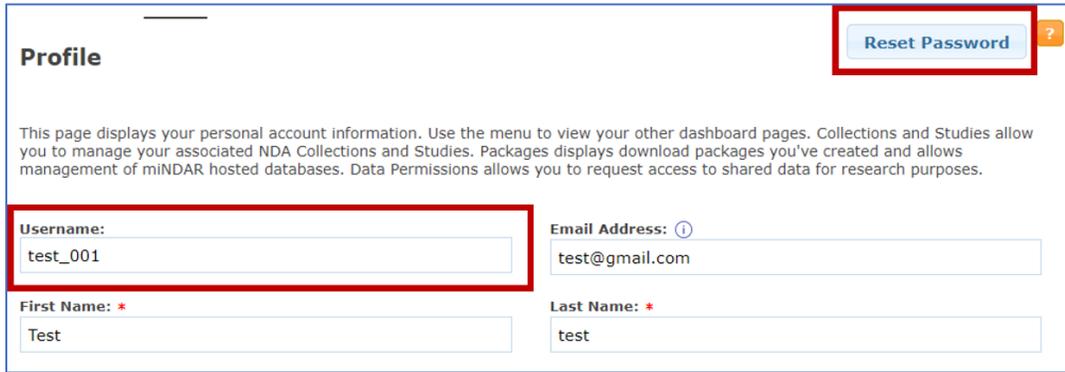


Figure 22: NDA Profile Username Field and Reset Password Button

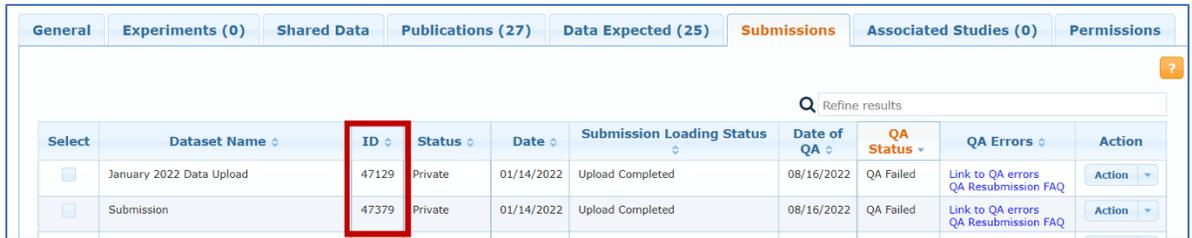
- b. Do **not** use your RAS identity credentials.
 - i. Your Login.gov, eRA Commons, or NIH credentials cannot be used with NDA-Tools.
- 3. Retrieve the submission ID containing QA errors you are ready to resubmit corrected data from one of the following methods:
 - a. The original submission confirmation message in the Python tool.

```
Uploads complete.
Checking Submission Status.

You have successfully completed uploading files for submission 52812 with status: Submitted_Prototype
```

Figure 23: Example of Submission ID returned by the Python tool

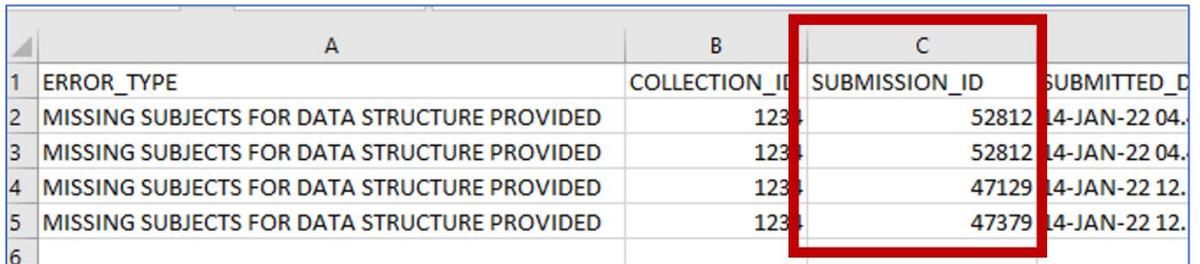
- b. The Submissions tab of your NDA Collection.



Select	Dataset Name	ID	Status	Date	Submission Loading Status	Date of QA	QA Status	QA Errors	Action
<input type="checkbox"/>	January 2022 Data Upload	47129	Private	01/14/2022	Upload Completed	08/16/2022	QA Failed	Link to QA errors QA Resubmission FAQ	Action
<input type="checkbox"/>	Submission	47379	Private	01/14/2022	Upload Completed	08/16/2022	QA Failed	Link to QA errors QA Resubmission FAQ	Action

Figure 24: Submission ID under Submission Tab of NDA Collection

- c. Column C in the .csv QA Results Error Report linked in the NDA QA results email notification.



	A	B	C	
1	ERROR_TYPE	COLLECTION_ID	SUBMISSION_ID	SUBMITTED_D
2	MISSING SUBJECTS FOR DATA STRUCTURE PROVIDED	1234	52812	14-JAN-22 04.
3	MISSING SUBJECTS FOR DATA STRUCTURE PROVIDED	1234	52812	14-JAN-22 04.
4	MISSING SUBJECTS FOR DATA STRUCTURE PROVIDED	1234	47129	14-JAN-22 12.
5	MISSING SUBJECTS FOR DATA STRUCTURE PROVIDED	1234	47379	14-JAN-22 12.
6				

Figure 25: Submission in Column C of QA Results Report

4. Validate your corrected data files by entering the following prompt in your terminal:

```
vtcmd -rs <submissionID> <correctedFiles.csv> -b
```

- a. Replace "submissionID" with the submission ID & "correctedFiles.csv" with the full file name and extension.
- b. Only upload files for the data structure that have need corrections.
- c. Do **NOT** resubmit any associated or manifest files.

```
C:\Users\alkohlanyrf\PycharmProjects\nda-tools-python\NDATools\clientscripts\vtcmd.py -rs
52812 ndar_subject01.csv -b
Running NDATools Version 0.2.19.dev17
-u/--username argument not provided. Using default value of 'alkohlanyrf' which was saved
in C:\Users\alkohlanyrf\NDATools\settings.cfg
 0%|          | 0/1 [00:00<?, ?it/s]
Validating files...
100%|#####| 1/1 [00:02<00:00, 2.01s/it]

Detected that all associated files have been previously uploaded from previous submission

Validation report output to: C:\Users\alkohlanyrf\NDA\nda-tools\vtcmd\validation_results
\validation_results_20221012T132253.csv

Note: Your data has warnings. To save warnings, run again with -w argument.
```

Figure 26: Example of Files Validating for Resubmission

5. Users will only have **one** chance to resubmit data using this tool. Before resubmitting, make sure:
 - a. All the QA errors are corrected.
 - b. All data structures that need correcting in your submissions are submitted at once.
 - c. All corrected data files have the correct number of rows.

6. You will have one last chance to confirm and resubmit with the prompt:
 - a. If your data files need to be corrected at this step due to one of the above, type "N" and press enter. This will cause the script to end.
 - b. If you've confirmed your data files are correct, type "Y" and press Enter.

```
All files have finished validating.

The following files passed validation:
UUID d42ac674-f727-4851-ae4c-8002e5c52377: ndar_subject01.csv
Building Package
Below is a summary of what your submission will look like with the validation files
provided:

Short-Name, Number of Rows
ndar_subject01,1
fmriresults01,40

Are you sure you want to continue? (y/N) y
```

Figure 27: Summary of Resubmission & Prompt to Continue

7. If you've followed the previous steps successfully, you'll receive a successful resubmission confirmation message:

```
Are you sure you want to continue? (y/N) y

Package Information:
validation results: ['b9cc4f33-100e-4a23-89c6-a818f6d1f5d2',
'd42ac674-f727-4851-ae4c-8002e5c52377']
submission_package_uuid: d9f3ea4d-d647-47f1-b2fe-a727e8dc85c7
created date: 2022-10-12T13:23:03.695-04:00
expiration date: 2022-10-13T13:23:03.695-04:00

Package finished building.

Downloading submission package.
Submission Package Download: 170kbytes [00:00, 295kbytes/s]

A copy of your submission package has been saved to:
C:\Users\alkohlanyrf\NDA\nda-tools\vtcmd\submission_package
\NDASubmission_alkohlanyrf_1665595384635
Requesting submission for package: d9f3ea4d-d647-47f1-b2fe-a727e8dc85c7
Submission ID: 52812

You have successfully replaced submission 52812.

Process finished with exit code 0
```

Figure 28: Successful Resubmission Confirmation Message

8. An unauthorized request message will appear if:
 - a. No QA errors were reported for the submissions ID provided, and resubmission was not necessary.
 - i. If you have not received a "QA Results" email notification from NDA requesting resubmission, no further action is required.
 - b. The wrong submission ID was provided.
 - i. Please check if you are using the correct Submission ID at Step 3 and try again.
 - c. The Incorrect NDA credentials are used.
 - i. Please check if you are using the correct Credentials at Step 1. If so, please check if you are connected to the internet and try again.

```
submission_id 52819 is not authorized to be replaced. Please contact the NDA help desk for approval to replace this submission
```

Figure 29: Unauthorized Request Message

NOTE: You can only resubmit their corrected data sets **once**. If you attempt to do so again, they will receive the following error:

```
Submission 52812 was already replaced by ndar_administrator on 2022-09-23T13:57:06.599-04:00. If you need to make further edits to this submission, please reach out the the NDA help desk
```

Figure 30: Resubmission Error - Submission Already Replaced

If you to edit an existing resubmission, please reach out to the NDA Help Desk at ndahelp@mail.nih.gov with you NDA Collection ID, submission ID and data structures you need to resubmit corrected data for.