BRICS USER GUIDE

MIPAV – Image Packaging Tool

CHAPTER 9 – MIPAV Image Packaging Tool

The Image Packaging Tool – also known as the Image Submission Package Creation Tool or the Imaging Tool – is a plug-in to the MIPAV (Medical Image Processing, Analysis, and Visualization) application that accepts the biomedical imaging data (in the form of **image file(s)** and, optionally, a **CSV file**) from a researcher and validates the metadata associated with the image files against the values defined in the Data Dictionary. It then creates an image submission package. The tool also provides a report of any data discrepancies, errors, and warnings received. If any validation errors are found, an image submission package cannot be created. In that case, the researcher should revise the data to fix all errors, first, and then re-validate the data.

The Image Packaging Tool is used to package and submit unprocessed and processed biomedical images in a variety of formats including DICOM, MINC 1.0 and 2.0, Analyze, NIfTI-1, AFNI and SPM. The tool also has the ability to process data organized as BIDS (Brain Imaging Data Structure) data. If you are using a different file format, please contact MIPAV support and/or BRICS Operations to have it added to the list of supported standards.

To submit imaging data to a BRICS instance's repository, you will need to use that instance's Submission Tools application with the submission package output by this tool. Using the Image Packaging Tool, you can prepare your image data for submission by following the steps outlined in the sections to follow.

The Image Packaging tool is a Java application and can be run on virtually any Java-enabled platform such as Windows, Linux, or macOS. The Image Packaging Tool is typically executed locally on your system as a Java Web Start application. In cases where the Web Start application cannot be run, please contact MIPAV support and/or BRICS Operations for assistance.

9.1 Objective

This chapter provides information for users on how to:

- Launch the Image Packaging Tool.
- Use the Image Packaging Tool manually.
- Use the Image Packaging Tool with an input CSV.

The process of validating and uploading imaging datasets to the Data Repository is covered in the User Guide on the Submission Tools application.

9.1.1 Supported Image Formats

The plug-in supports all file formats supported by MIPAV including processed and unprocessed medical images including the DICOM and NIfTI formats. The plug-in can also process data organized in the BIDS format. If you are using different file format not supported by MIPAV, please contact MIPAV support and/or BRICS Operations.

9.1.2 System Requirements

The user must have Java Runtime Environment (JRE) version 8-10 to use the Image Packaging Tool. Users with JRE versions 11 or newer may not be able to run the application using Java Web Start. If this happens, please contact MIPAV and/or BRICS Operations for support.

Use this link to download JRE version 10: 64-bit Java Runtime Environment (JRE).

9.1.3 Preparing Image Data for Submission

Using the Image Packaging Creation Tool, you can prepare your image data for submission by following the steps outlined in the next sections. When medical image files are loaded into the Image Packaging Tool, the tool extracts out any available image header metadata, and attempts to map the image header metadata onto the Data Elements in the selected Form Structure. The quality and amount of image header metadata that can be extracted out of an image volume will depend on the medical image file format, the scanner on which the images were acquired, and the de-identification process performed. There are two ways in which you can load imaging data to the BRICS repository using the Image Packaging tool:

1. **Manual Process:** If you are loading images manually, you may use the "**Add Record"** button to manually choose a BRICS Form Structure and enter data element values for that Form Structure. Follow the manual loading process outlined in the sections below to load images.
2. **Batch Process:** If you are loading multiple images at a time, follow this process to fill out a CSV template provided to load your images without manual data entry via the Image Packaging Tool.

Module Input
1. **De-identified Medical Image(s)** in one of the supported formats.
2. **Corresponding Input CSV** files with metadata, if using the Batch Process.
3. **Additional Metadata** describing the medical images not available via the image file header, or the user's input CSV file(s).
   a. **Some fields, such as the subject's GUID, cannot be automatically extracted from the imaging data. By providing this information in the input CSV, the user can avoid having to manually input this information within the Image Packaging Tool.**

Module Output
1. The original medical image(s) in one of the supported formats (zipped up if the image volume is comprised of more than one file on disk).
2. Output CSV file(s) for each Form Structure in your submission, ready for validation by the Data Validation module.
3. An image thumbnail file (JPEG) for each medical image volume.

4. The Output log that lists all image files and CSV files added to the image submission package. It also displays the path(s) to the directory where the image package(s) is stored.

9.1.4 Image Packaging Tool Input

Data can be loaded into the Imaging tool in the form of medical image file(s) and optional input CSV files that contain additional patient/subject/visit information (not stored in the image header) as well as image related metadata.
The following information is required for each primary image file in an image submission:
1. The GUID for each patient/subject.
2. Certain image-specific information: the imaging file(s), some image acquisition metadata, image QA/QC information.

Medical Image File Format Support
The Image Submission Packaging Tool supports dozens of medical image file formats, including NIfTI, and many variants of the DICOM format.

When adding your image data to the Image Submission Packaging Tool, you have different options for how to select and package your de-identified data files:

-
- A ZIP archive containing the files comprising the image dataset (e.g., all of the slice files of a multifile DICOM image volume, or the .img/.hdr file set of an Analyze format dataset) and only the files from the image dataset.

- A gzipped tar archive (commonly referred to as a tarball, or .tar.gz format), containing all file files of the image dataset.
- If the image file format of your data is a single-file format (such as a .nii NIfTI file, or a multi-frame DICOM file), you can directly select the image file via the Imaging Tool graphical user interface, or through an input CSV file.
- If the image file format of your data is a multi-file format (such as most DICOM datasets, where each slice of the image is stored in a separate file), you can load the dataset in the Imaging Tool user interface by checking the "Open as multifile" checkbox in the "**Image Information.ImgFile**" Browse dialog before selecting any file from the , or when using an input CSV by specifying a directory containing only the image volume files (from only series/acquisition) in each row of the "**Image Information.ImgFile**" column.
    - Note that all of the individual files must be in the same directory and that directory must not contain any other imaging files that are not part of that image series.

9.1.5 Output Image Packaging Tool

The imaging data submission package includes:

- A compressed zip file for each image series containing the selected image files, or the uncompressed image file for each record, if using a single-file format.
- A JPG file for each primary image series that can be used as a thumbnail to preview the image in the Query Tool.

- A CSV file with the metadata describing the image(s).

The metadata CSV file and a subdirectory containing the image file(s) and JPG thumbnails will be created in the Output Directory selected in the Image Packing Tool.

9.2 Downloading and launching the Image Packaging Tool

To submit imaging data to the BRICS repository, you are required to run a Java Web Start application, to properly prepare your data for Validation and Upload.

For using the Image Packaging Tool, you can prepare your image data for submission by following the steps outlined in the next sections.

To download and launch the Image Packaging Tool follow these steps:
1. Log into your BRICS instance and Data Repository Module:
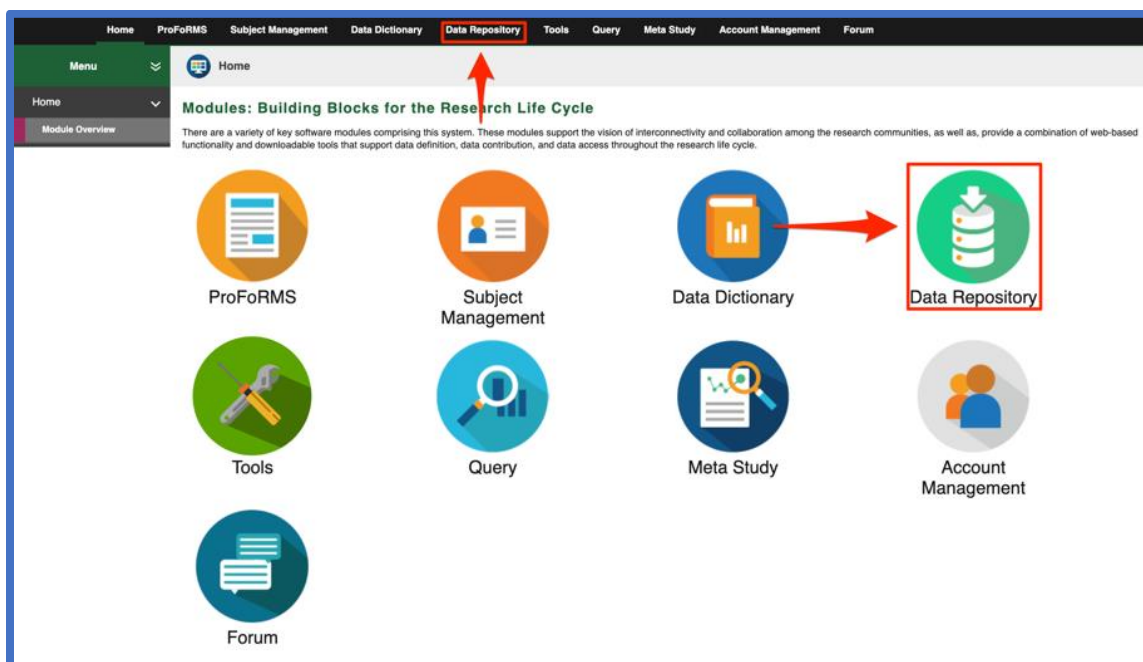
2. Expand Submission Tools in the left navigation and then select Image Packaging Tool:

3. Here select Launch Image Packaging Tool and then either save the tool to your computer or run it immediately.

4. To continue, you **MUST** read and accept the warning banner by Clicking on the **Yes** button to run and retrieve imaging structures from BRICS Data Dictionary. Click the **No** button to exit the application.

5. The Image Packaging Tool should open:
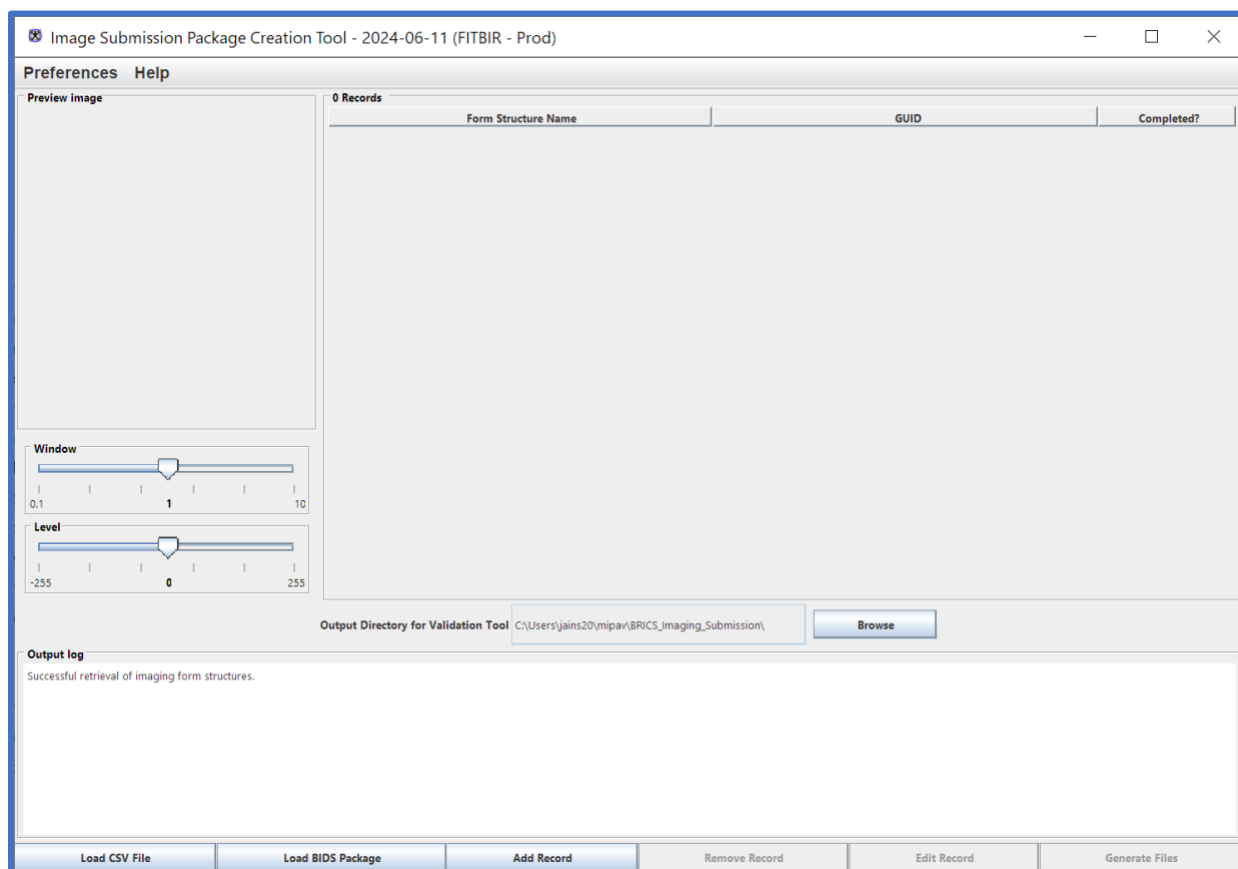
9.3 Using the Image Packaging Tool

Image Submission Packaging Tool/MIPAV Data Submission Workflow:
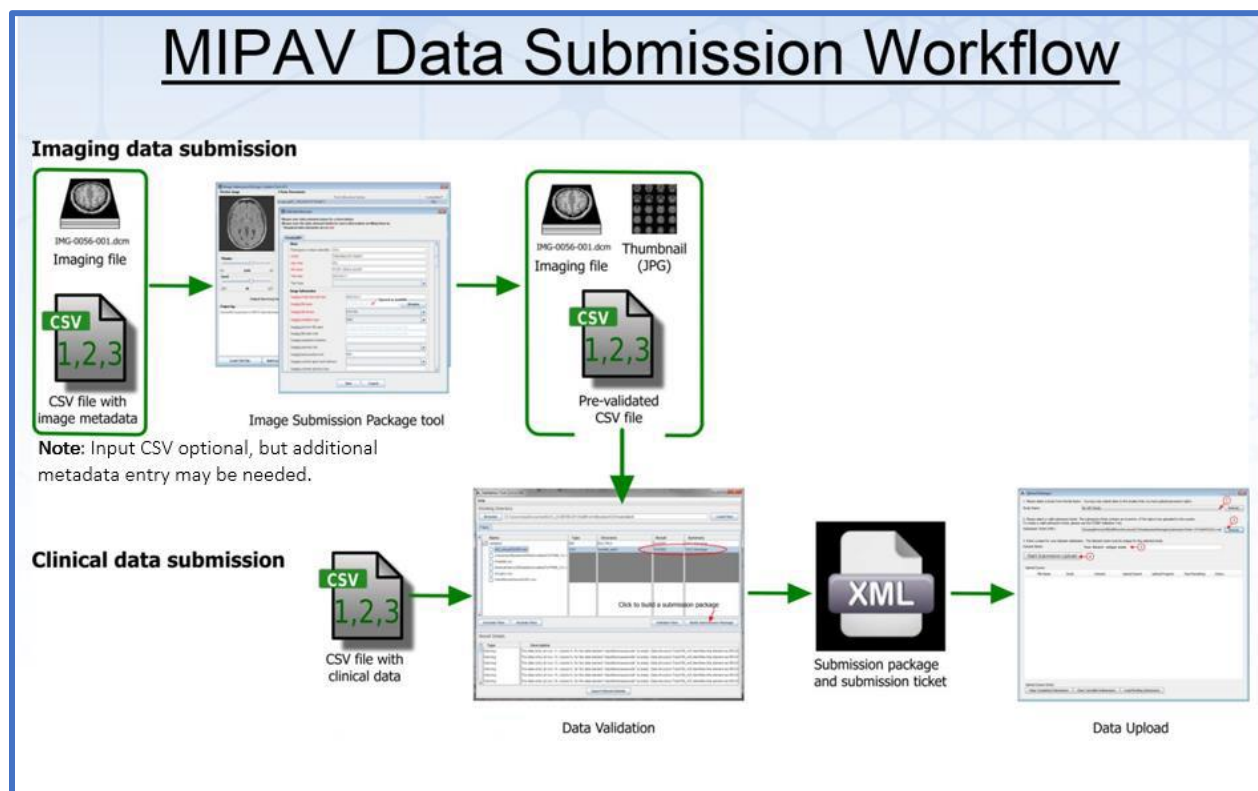
For Manual Loading Process Perform the following actions:
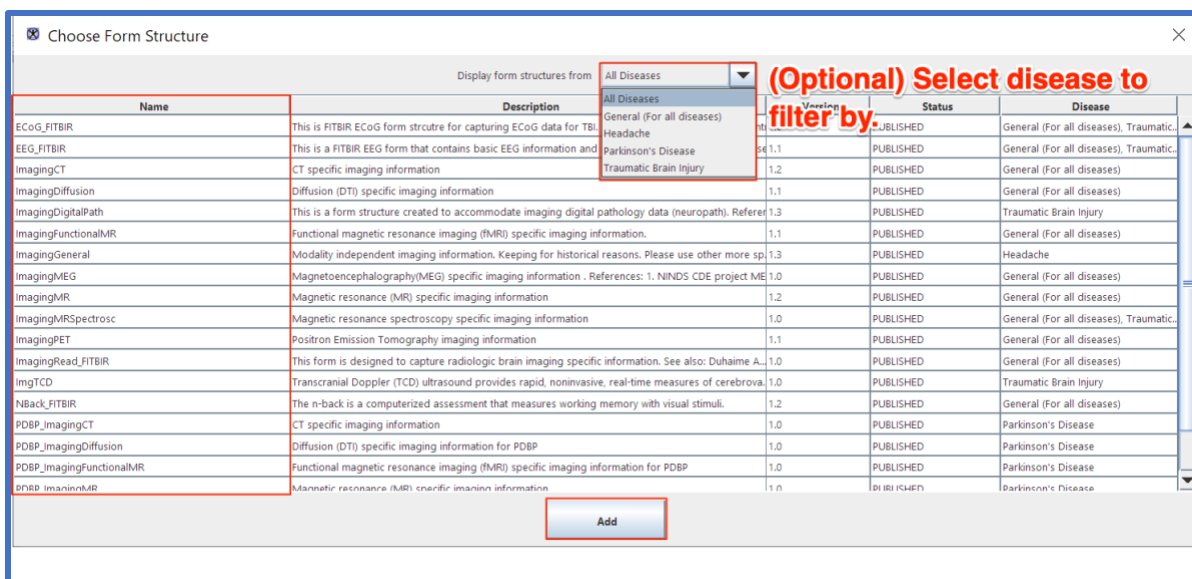
1. Launch the Image Packaging Tool by following the instructions as outlined in Section 9.2 of this guide.
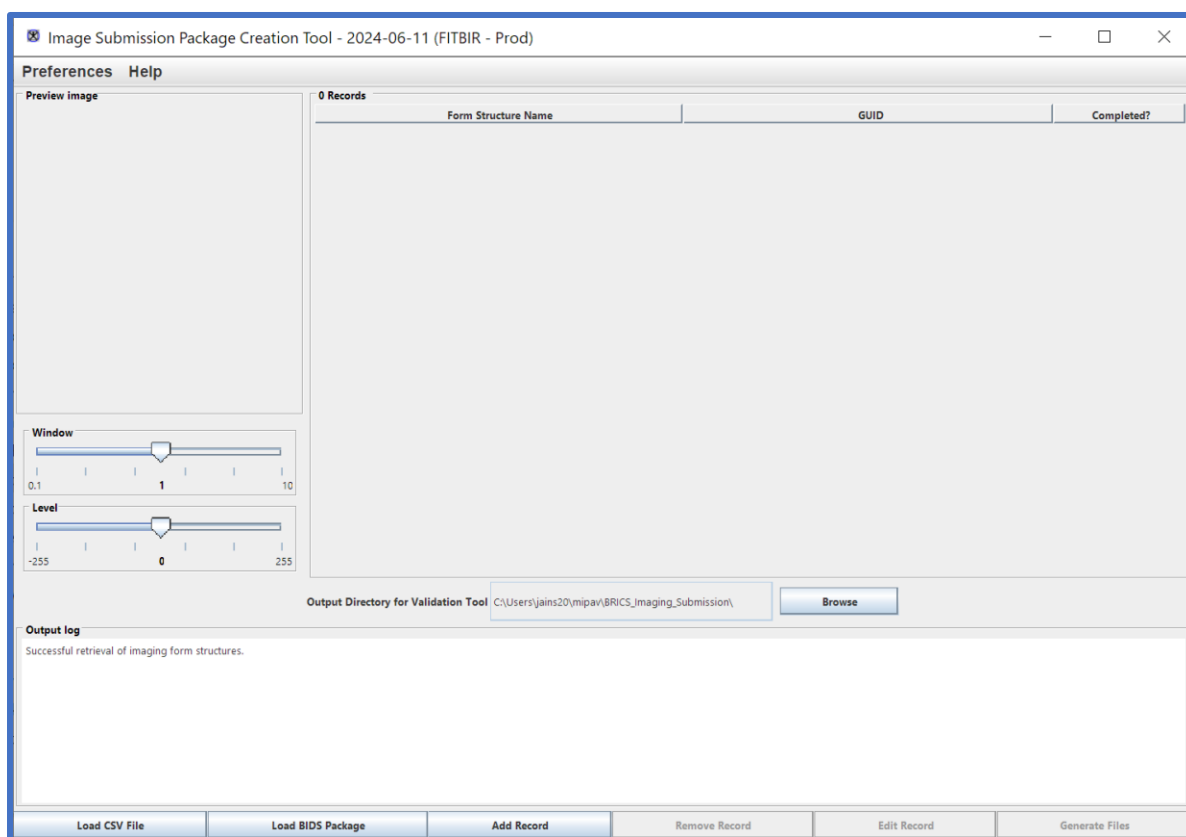
2. In the main window, Click the **Add Record** button to choose a published Imaging Form Structure from the BRICS Data Dictionary.

National Institute of Health/CIT – Proprietary and Confidential
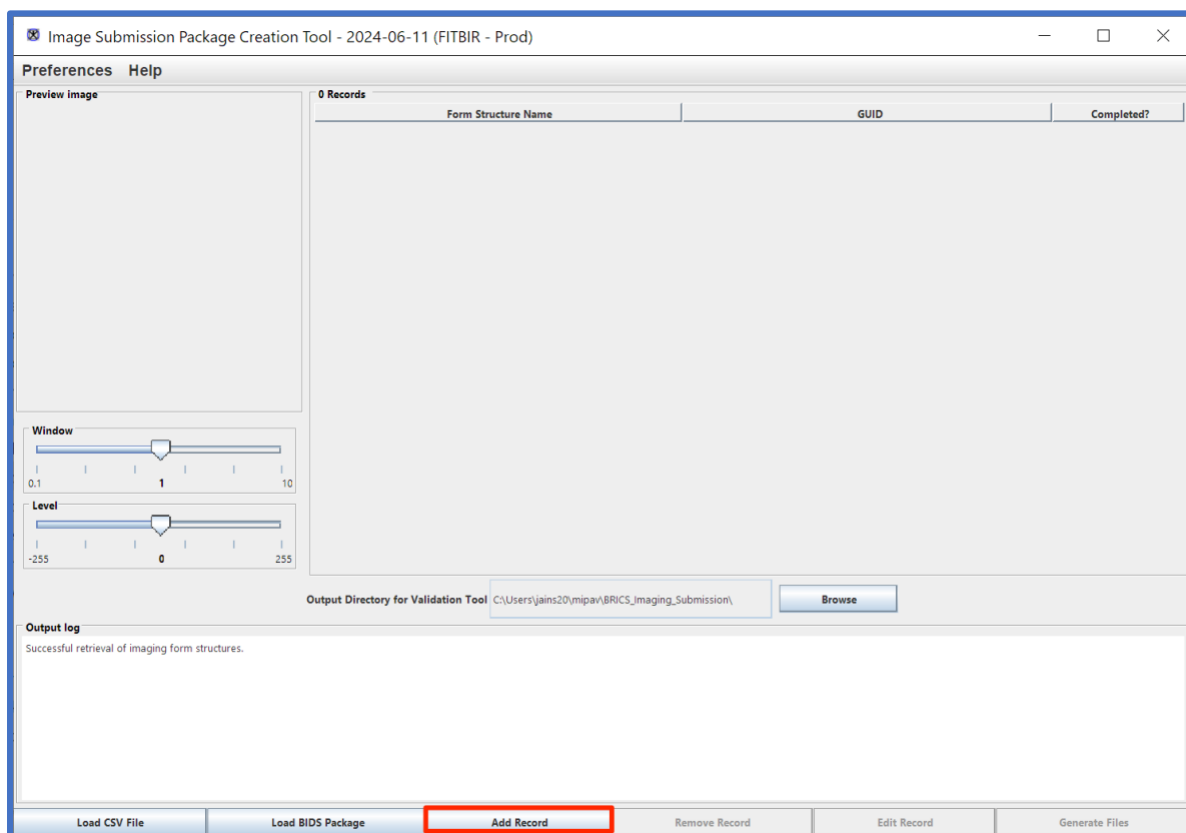
3. The **Add Record** dialog box appears, listing the form structures available from the BRICS data dictionary. You can filter the Form Structures on Disease by selecting from the drop-down menu at the top. Select the Form Structure that is appropriate for the image volume you want to load next and Click the **Add** button.

4. The **Edit Record** window appears populated with i) the data elements from the selected form structure and ii) with metadata from the CSV file. Review the information in the Edit Data Elements window. Make sure that all required fields are filled in.

5. To load the image data into the Image Submission Package Creation Tool, navigate to the **Imaging File** field and use the **Browse** button to open the Open File dialog box and select the file.
    The **Edit Data Elements** window appears populated with the data elements from the selected form structure. Review this information. The required fields appear in **RED**. You can place your mouse over the Data Element names on the left and the Data Element fields on the right to see more information about each Data Element and any guidelines for its entry.

6. In the dialog that appears, browse to the location of the image acquisition you want to load into the Imaging Tool. If your image dataset is contained within a single file (for example, it is in single file NIfTI format, a single file DICOM, or a .zip/.tar.gz file containing all the image files comprising your image series), select the file and press the Open button. If your data is stored in multiple files on disk (for example, it is in one slice per file DICOM format and has not been put in a .zip file), select any one file from the DICOM dataset, check the "Open as multifile" button on the dialog, and then press the Open button. Checking the "Open as multifile" button will instruct the Imaging Tool to search the parent directory of the file you selected for other DICOM files that are part of the same DICOM series. Once the Imaging Tool finds all these files, it will load them all together as one 3D or 4D image volume, as appropriate.

National Institute of Health/CIT – Proprietary and Confidential
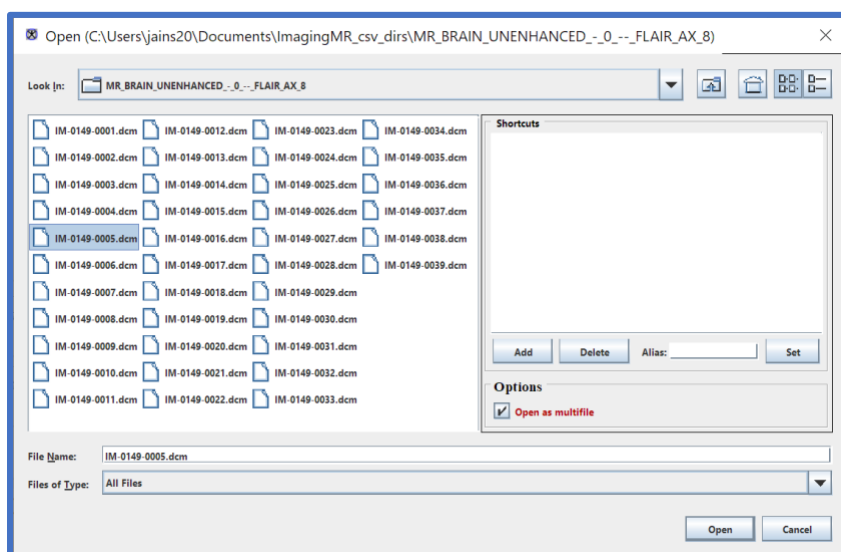
7. If your image data files contain any DICOM header tags that may contain personally identifiable information, a new dialog will appear, showing you the potentially problematic fields and their values. In the **De-Identification Review** dialog window, review all the fields that may contain PII/PHI. Click on the "I have reviewed the data and no PII/PHI is present" to continue OR click the "Exit the Imaging Tool" button to leave the tool and perform additional de-identification on your image data.

National Institute of Health/CIT – Proprietary and Confidential

8. After you have selected an image dataset, it will be loaded into the Imaging Tool, and the Imaging Tool will attempt to map any pertinent image header metadata from the files to the Data Elements in the current Form Structure. The quantity and quality of the extracted image header metadata will vary depending on the image file format (with DICOM generally providing the most information), scanner manufacturer, and any post-acquisition processing that was performed. If you have selected a NIfTI format file, the Imaging Tool will attempt to read a JSON file from the same directory as the NIfTI file, replacing the .nii or .nii.gz file extension with .json. If a matching JSON file is found, the Imaging Tool will map metadata from the JSON file onto the Form Structure Data Elements. Review the extracted Data Element values, and then manually enter values for required Data Elements and any missing image information that is important for fully describing your data to other researchers.

9. The required fields appear in red. If some of required fields are empty, the image submission package will fail validation.

10. Review the information entered in the previous step. Make sure that all required fields (appear in red) are filled in. If some of required fields are empty, the form structure status will appear as not completed and you will not be able to generate an image submission package for later validation.

National Institute of Health/CIT – Proprietary and Confidential

11. After all the Data Element values necessary to fully describe your imaging data for use by other researchers (not only the set of required fields) have been filled in, either via image header data extraction or man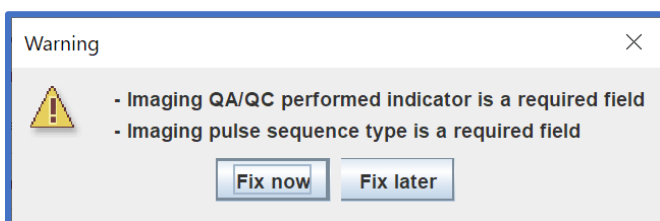ual entry, click the **Save** button. **Note** that depending on the form structure, the required data elements may vary.

12. If any required fields were not completed, you will be prompted to decide whether to fix them now or defer them.

a. Choosing to "Fix now" will bring up a dialog highlighting the incomplete required Data Elements. Enter the missing image information until all the required fields are complete. You can scroll through any incomplete Data Elements via the Prev, Current, and Next error buttons.

b. Choosing to "Fix Later" will bring up this showing the record marked as incomplete. Double clicking on the record from the main screen will allow you to edit the record again.
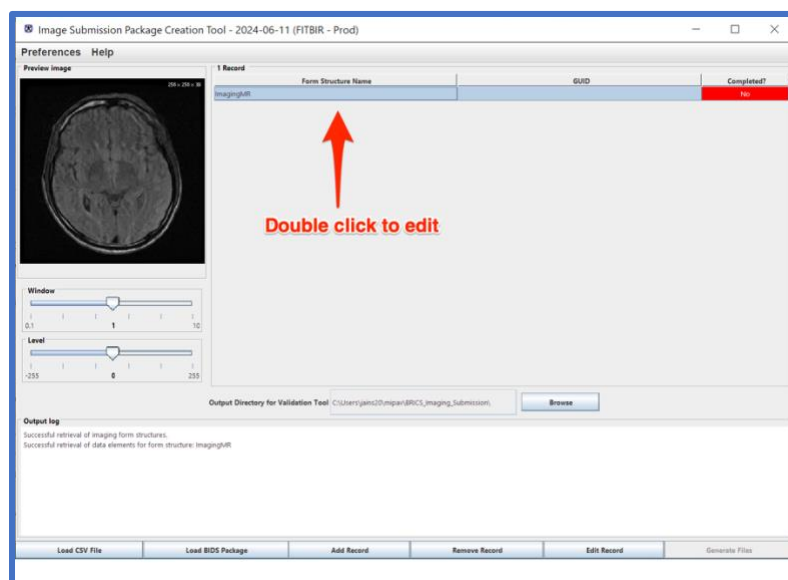
13. Specify the **Output Directory** which would be used by the module to store temporary files, validation logs and image submission packages.
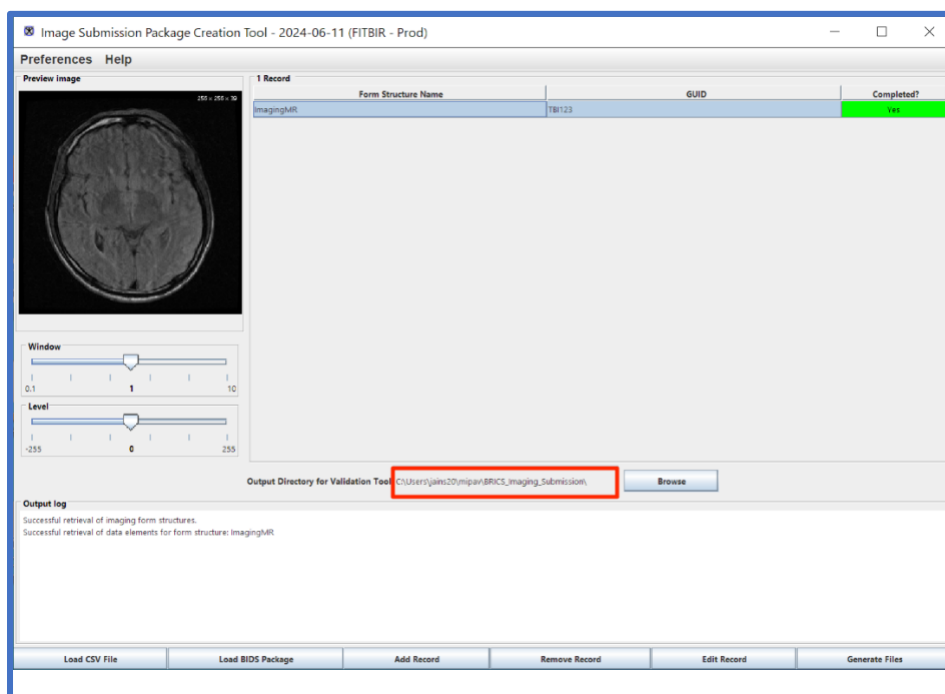
14. Click the **Generate Files** button. This will generate the image submission package. The output log message appears in the Output log window showing the progress, the image submission package file name(s) and location, and other helpful information. Finally, the image submission package appears in the **Output directory**.

National Institute of Health/CIT – Proprietary and Confidential
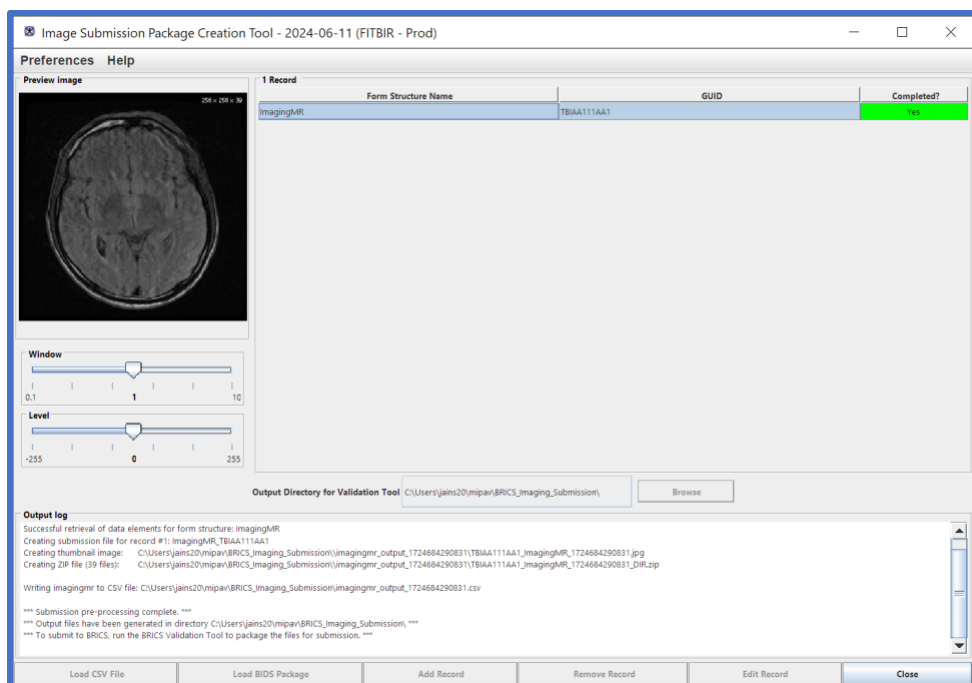
9.3.1 Creating a CSV Template
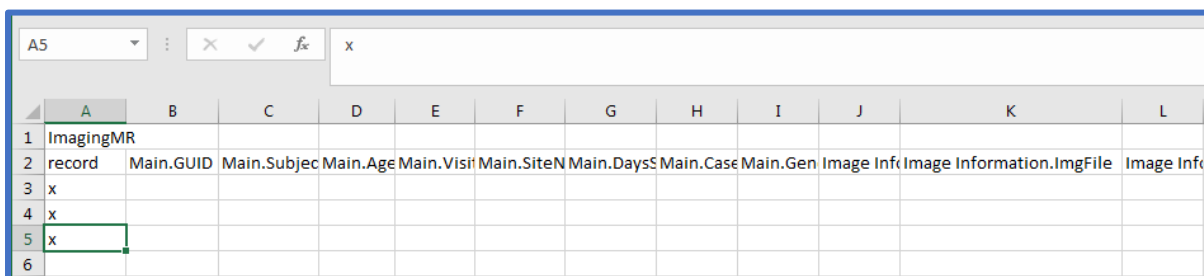
To create a CSV template: Perform the following actions:

1. Navigate to your Workspace and click the **Data Dictionary** module.
2. Select your **Form Structure** from the list of form structures.

3. Click the **Export Form Structure** and choose CSV to export or Select CSV with sample data to export.
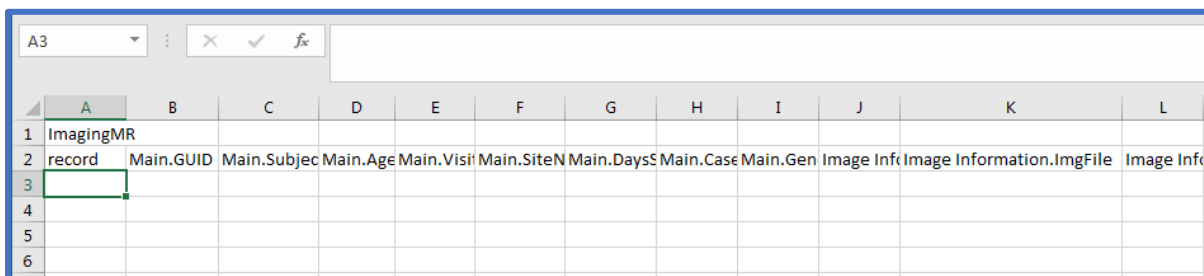4. Select the radio button **Open with Microsoft Excel (default)** to open the CSV template Click the **Ok** button.

5. The CSV template will contain the name of the Form Structure and a listing of Data Elements organized by their Repeatable Group.

6. For each imaging record, place an 'x' in the "record" column.

    a. Note that some imaging records use multiple rows (e.g. when there are multiple auxiliary files for ImagingFunctionalMR image series). The record column is used to indicate the start of a new record.

7. Provide important Data Element values for each imaging record. Pay special attention to the "Imaging Information.ImgFile" column, where you can specify the image data file to load for each imaging record. This imaging file can be a single file medical image file format (such as a .nii NIfTI file), a .zip or .tar.gz file containing a multi-file medical image acquisition (such as a multi-file DICOM image volume/series), or the path to a directory containing only the files constituting a multi-file medical image acquisition. When your input CSV is read by the Image Packaging Tool, the file or files indicated in each row of this Data Element will be read in, and metadata in the image header will be mapped on to the other Data Elements for this image record.

8. Fill in the values for any other Data Elements that are necessary to fully describe your data. Any Data Elements whose values will be provided via the image header data mapping do not need to be specified in the input CSV. Aside from the "record" column, the CSV columns may be reordered, or removed if they are empty.

9. Continue steps 7 through 9 again for any other image data you want to include in this input CSV. Each input CSV should contain a maximum of approximately 100 image records. If you encounter an out of memory error while loading an input CSV, please contact MIPAV **support** and/or BRICS Operations.

   a. The reason for this maximum is to make sure that the resulting image submission package is not too large. Larger imaging datasets are often more difficult to upload and to manage on BRICS instances.

10. Save your input CSV for use later in the **Image Packaging Tool.**

9.3.2 Using a CSV Template

Follow these steps to use a CSV template:
1. Launch the Image Packaging Tool by following the instructions as outlined in Section 9.2 of this guide.
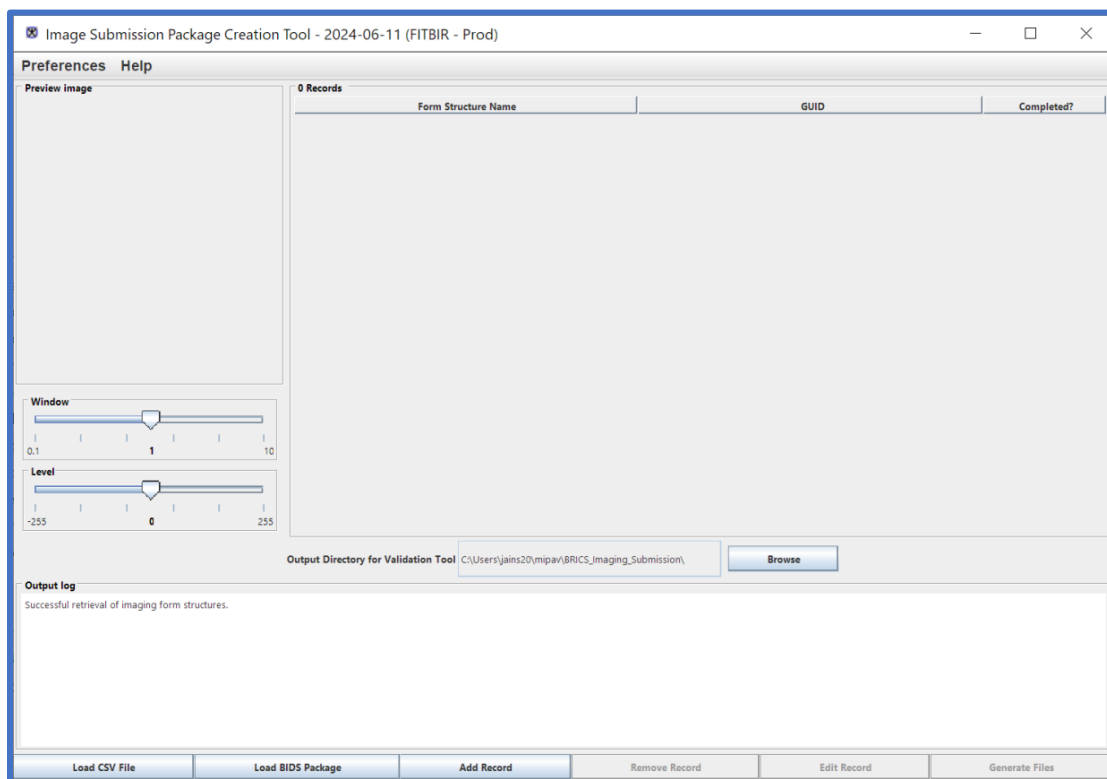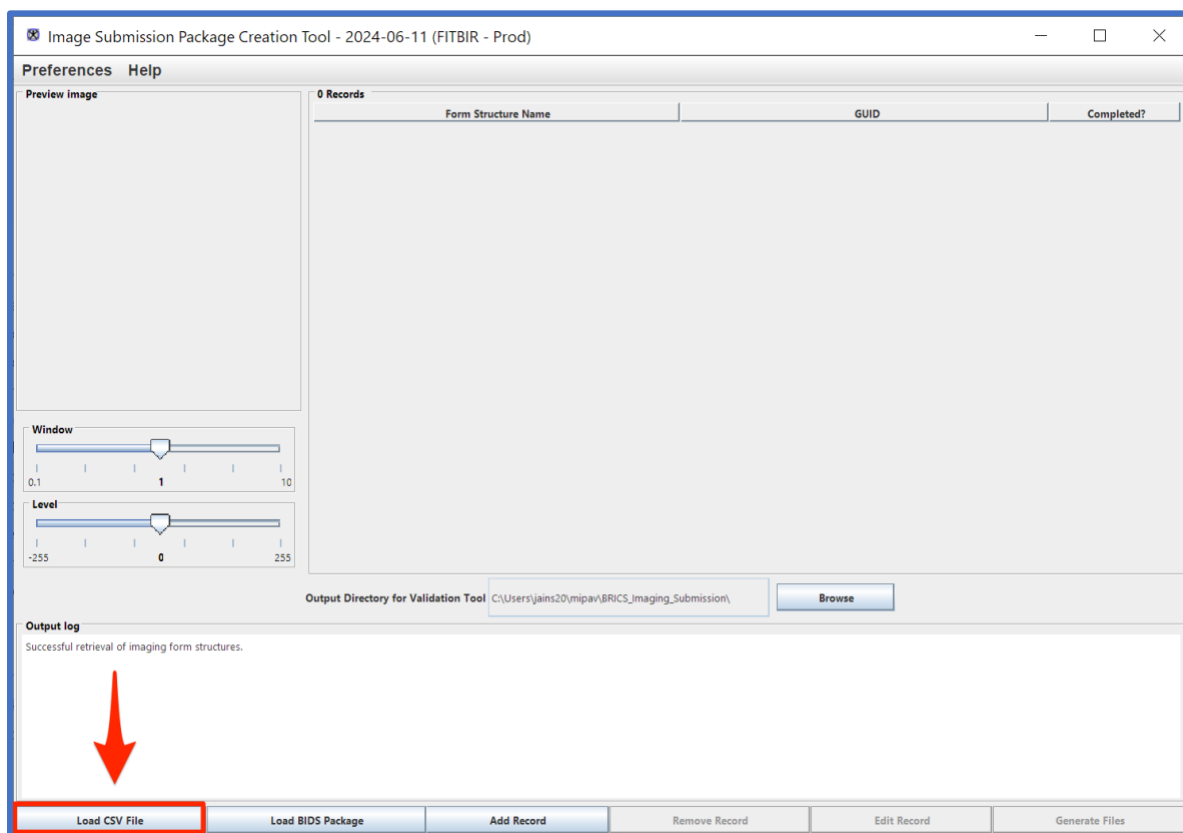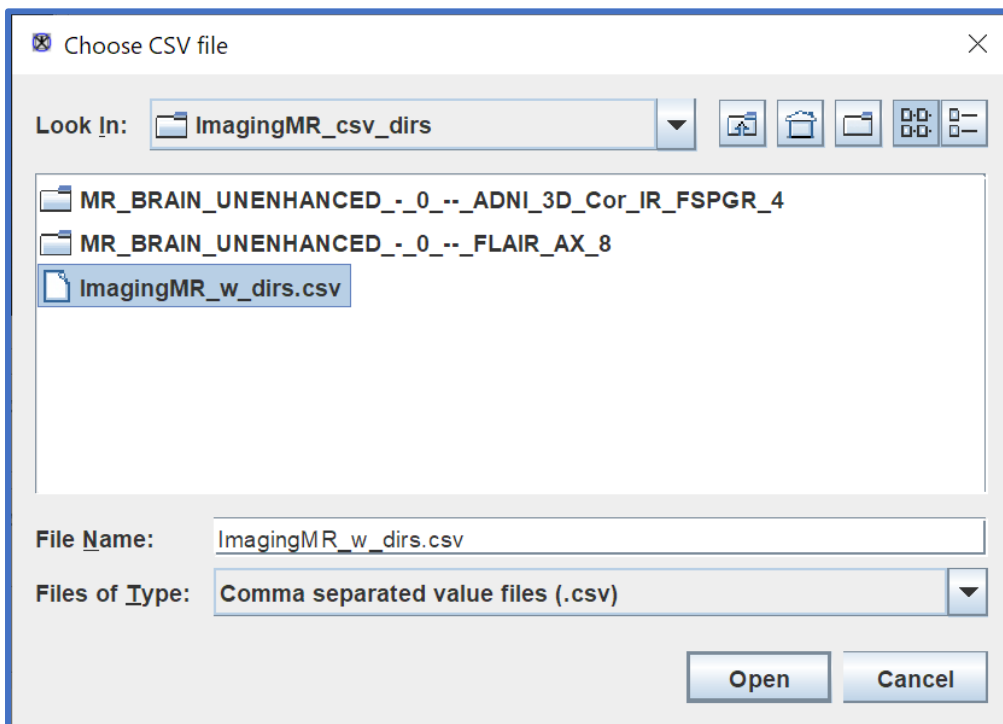
2. In the main window, use the **Load CSV File** button to add the image file(s) and metadata file(s) (in CSV format) needed to create the image submission package.

3. In the **Choose CSV file** window that appears, use the navigation buttons to navigate to the folder where you store your CSV files, select the CSV file for upload and Click **Open**.

4. The Imaging Tool will read in your selected CSV and load the imaging data referenced in each row of the CSV. Header data from each image dataset will be mapped onto the Data Elements of the form structure used by the input CSV and will be merged with any Data Element values specified on that same row of the input CSV. If a conflict exists between the value specified in the input CSV and the value extracted from the image header, the user will be shown each value and prompted to select which of the two values should be used in the output CSV metadata as shown below:
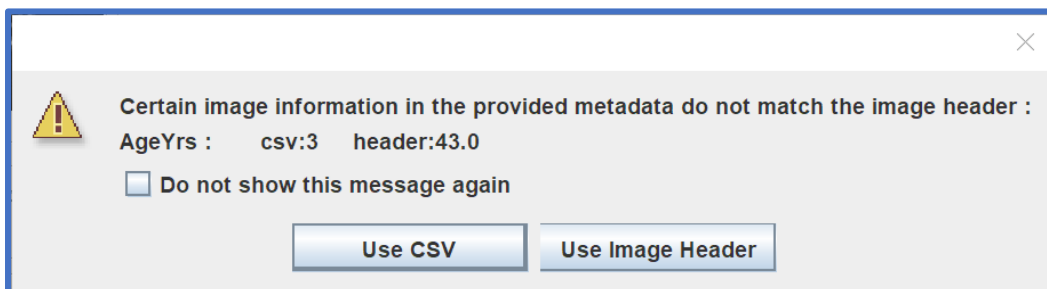
Checking this box will use the selected option for all future discrepancies encountered while reading in the data specified in this CSV.

5. After all the rows in the input CSV have been processed, A **De-Identification Review** dialog window will appear for each image volume where values were discovered in potentially PII/PHI-related header locations. Review all the fields that may contain PII/PHI. Click on the "**I have reviewed the data and no PII/PHI is present**" to continue OR click "**Exit Imaging tool to discontinue the uploading process**".

De-identification review: IM-0149-0001.dcm

The table below lists fields in the loaded image data with potential Personally Identifiable Information (PII) or Protected Health Information.

Please review all the fields below. If any fields contain PII/PHI, exit the Imaging Tool and fully de-identify your image data.

There may be fields in your data that contain PII/PHI that are not highlighted in this table. DICOM private tags, and sequence tags are not examined.

Remember, *YOU* are responible for the de-identification of all submitted data. This table is for informational purposes only.

Base file loaded:
C:\Users\jains20\Documents\ImagingMR_csv_dirs\MR_BRAIN_UNENHANCED_-_0_--_FLAIR_AX_8\IM-0149-0001.dcm

| DICOM Tag	Name	Value
0010,0010	Patient's Name	173185
0010,0020	Patient's ID	SF-1114

| I have reviewed the data and no PII/PHI is present | Exit the Imaging Tool |

6. The **CSV file** appears in the main window in the **Form Structure Name** table. The corresponding image file appears in the Preview Image box.
If the image file did not appear in the **Preview Image box**, review the error message issued by the system. Chances are that the CSV file has a wrong image file name/location. Review the CSV file and fix the image file name.
In some cases the CSV appears as not completed (Status Completed is set to No). This means that some essential information is missing, and you need to add this information manually.

National Institute of Health/CIT – Proprietary and Confidential
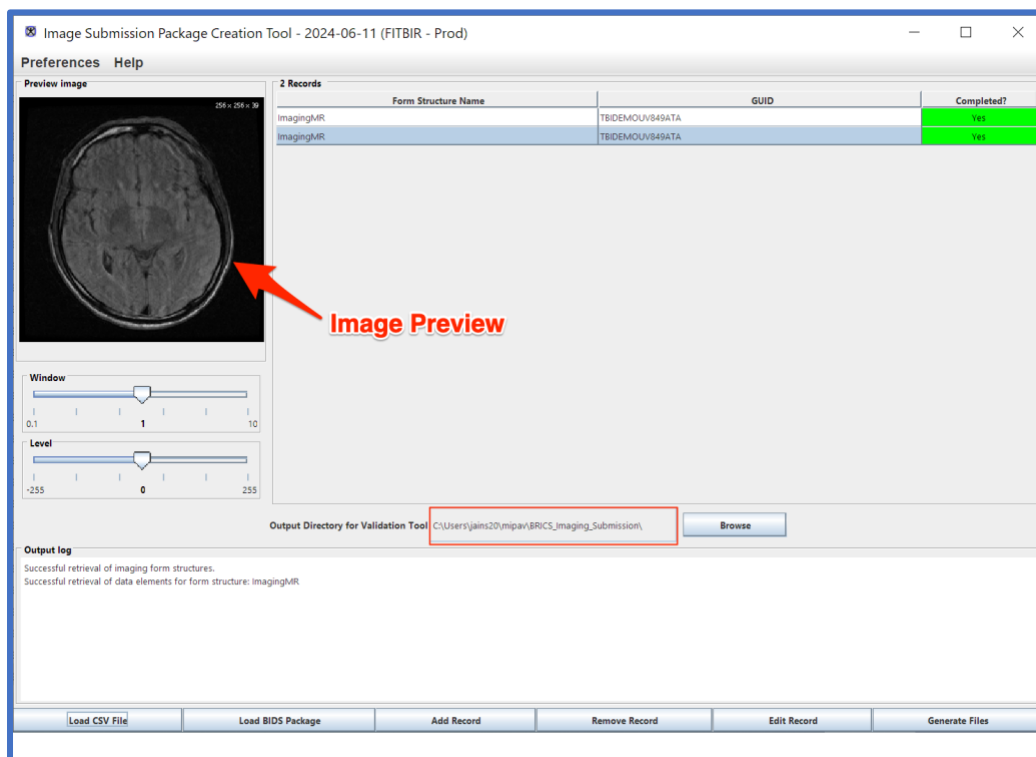
7. In the main module window, select the **Output Directory** which would be used by the Imaging Data Submission and Validation tool to store temporary files, validation logs and image submission packages.
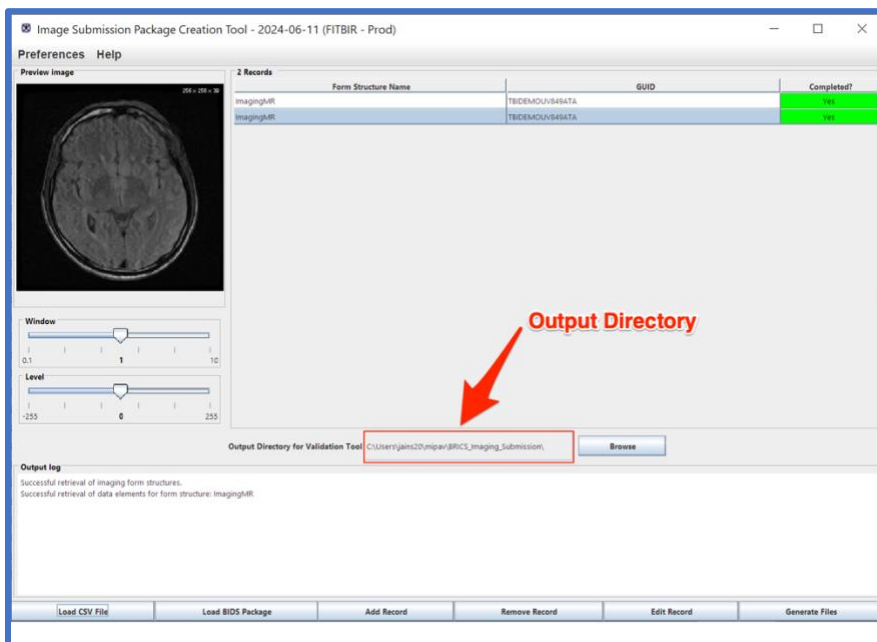
8. Click **Generate Files**. This will generate the image submission package.
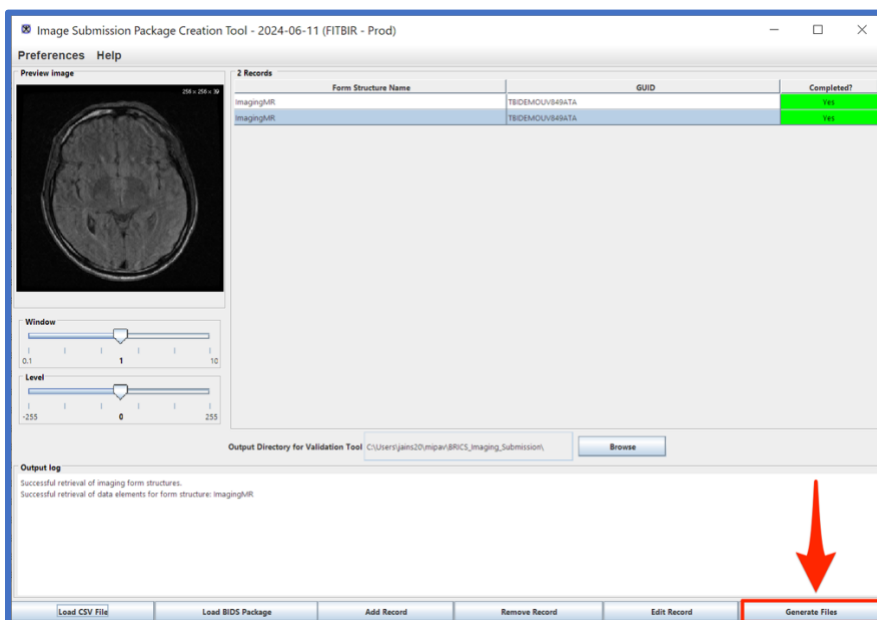
9. The **Output log** message appears in the **Output log** window showing the progress, the image submission package file name(s) and location, and other helpful information. Finally, the image submission package appears in the Output directory. Click the **Close** button to close the **Image Packaging Tool**.

     National Institute of Health/CIT – Proprietary and Confidential