



File Checker User's Guide

Introduction

To enable their research and quality improvement missions, the Multicenter Perioperative Outcomes Group (MPOG) provides its members with a software framework for loading and managing their clinical data. This framework is called Import Manager. Data is loaded into Import Manager through the use of flat text files. In order to be accepted, these files must adhere to a strict specification. This specification defines the acceptable structure and formatting for the files as well as the schedule on which they are to be generated.

File Checker is a development tool to help sites verify that the files they create for MPOG will be accepted. This application runs sample files against a battery of tests, each checking a specific part of the specification. Users are then informed of any defects a file may possess, including the specific row whenever possible.

File Checker is a stand-alone application that does not require Import Manager or any other MPOG software to be installed. This allows prospective members to try out developing file extracts without investing in costly hardware.

For more information on how to get started generating files for this application, please refer to the file specification and Import Manager playbook documents (included separately).

Disclaimer: File Checker is only a first pass of validation

While it will detect many potential issues, File Checker cannot find or predict every problem that may arise. As such, files that pass all tests are not guaranteed to be accepted by the Import Manager framework. With that said, using this application will drastically reduce development time by allowing quicker iteration.

Who are You?

There are two main types of users of File Checker: those responsible for all files and those responsible for a subset of files. Which type you are can affect how you will use this application, as explained below.

I am responsible for all files imported into the MPOG framework

Examples: Main extract developer, technical lead for MPOG go-live project

As data files may come from a variety of sources, each extract developer typically can only validate their subset of files. However, many tests within File Checker cross-reference identifiers to ensure that data between files can be properly linked. Therefore, it is important to collect and test a complete set of files before loading them into Import Manager. You will be responsible for verifying all files, even if you do not create all or even any of the files yourself.

I am only responsible for creating a subset of files

Examples: third-party billing vendors, report writers for supplemental extracts

If you were asked to develop only a subset of file extracts, you will likely not have access to the full set of files during your testing. Therefore, certain checks within the application will not be available to you (e.g. whether identifiers between files match). However, File Checker will still be able to verify other elements of the specification such as number of columns and data types. Also, you can safely ignore any failures from the “required file types are included” test.

How to Use

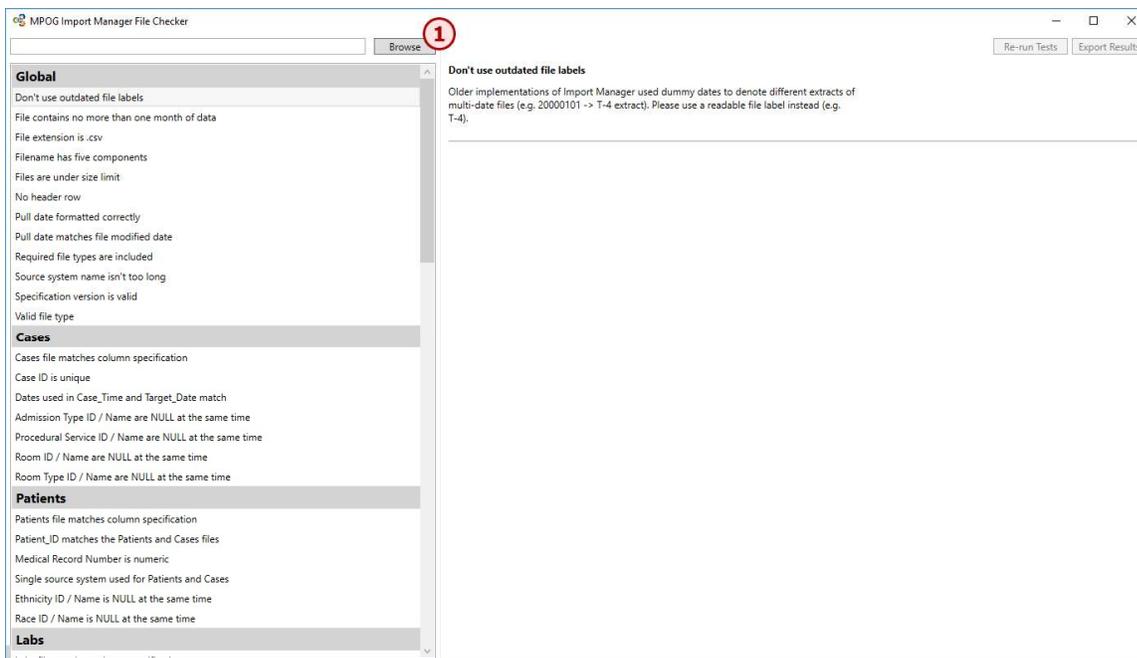
Setup

File Checker is a stand-alone executable that does not need to be installed. Simply double-click on the File Checker icon to open it. When the application opens, File Checker will appear blank as no files have been loaded. To begin, you will need to load one or more for testing.

Loading Files for Testing

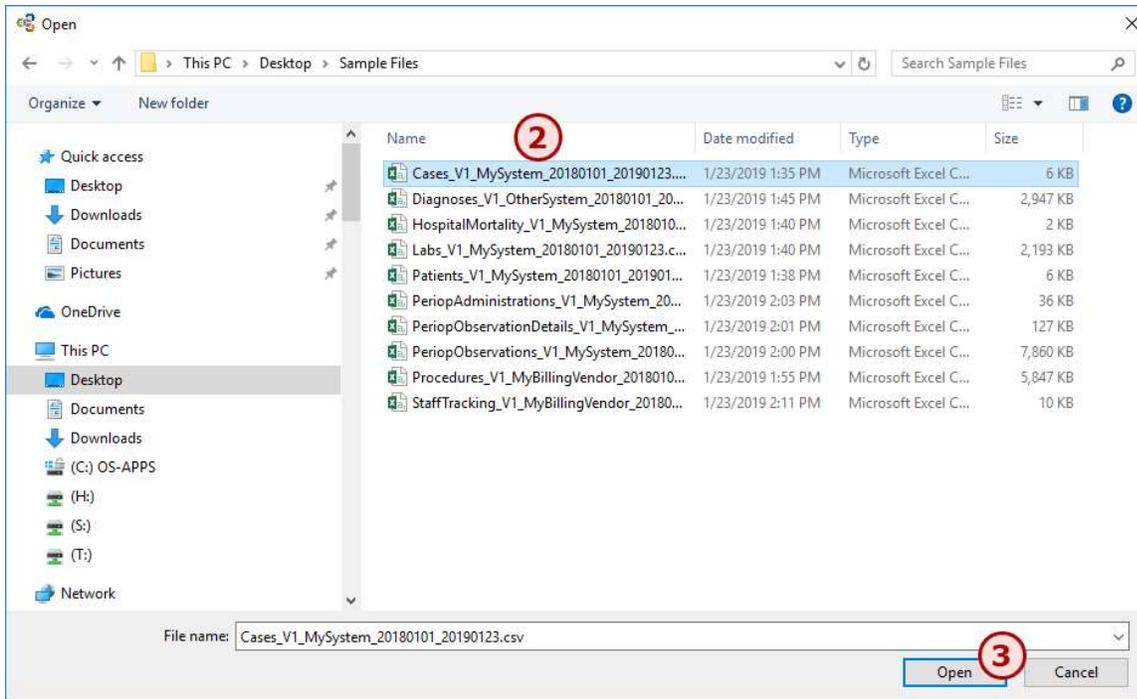
In order to load in one or more files, do the following:

1. Click the Browse button.



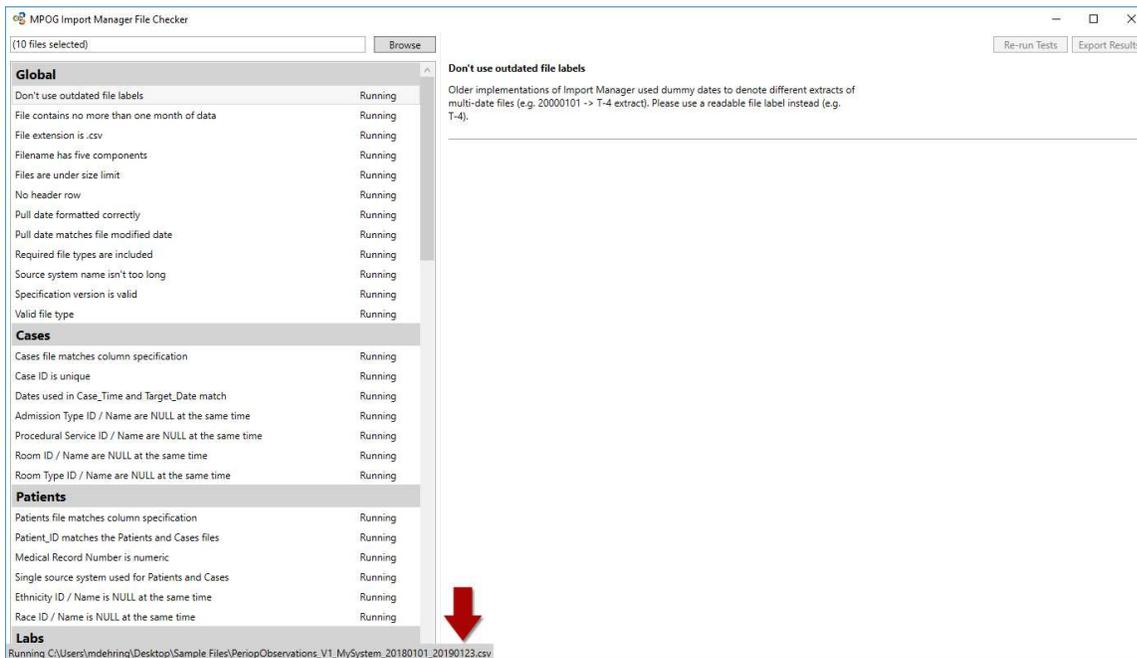
Browse button to load file(s)

- In the open file dialog, navigate to the folder containing the data files for MPOG. Select the file(s) you wish to test.



Open File Dialog

- Click the Open button. The application will immediately begin scanning files.



The file currently being scanned is indicated in the bottom-left corner.

Note that you can either test one file at a time, or groups of files at once. Checking one file can be useful for quicker iteration for a particular extract. However, it is recommended to test groups of files whenever possible as File Checker can cross-check identifiers between files (e.g. a Case ID used in PeriopObservations is the same Case ID used in Cases). When testing multiple files at once, make sure that all files contain the same date range.

The Open File Dialog uses standard Windows file selection methods, such as:

- Control-click to add/remove individual files from the current selection
- Click then shift-click to select a contiguous range
- Click and drag to rectangle select
- Keyboard shortcut CTRL-A to select all files in the current folder

Warning: Do not load large files

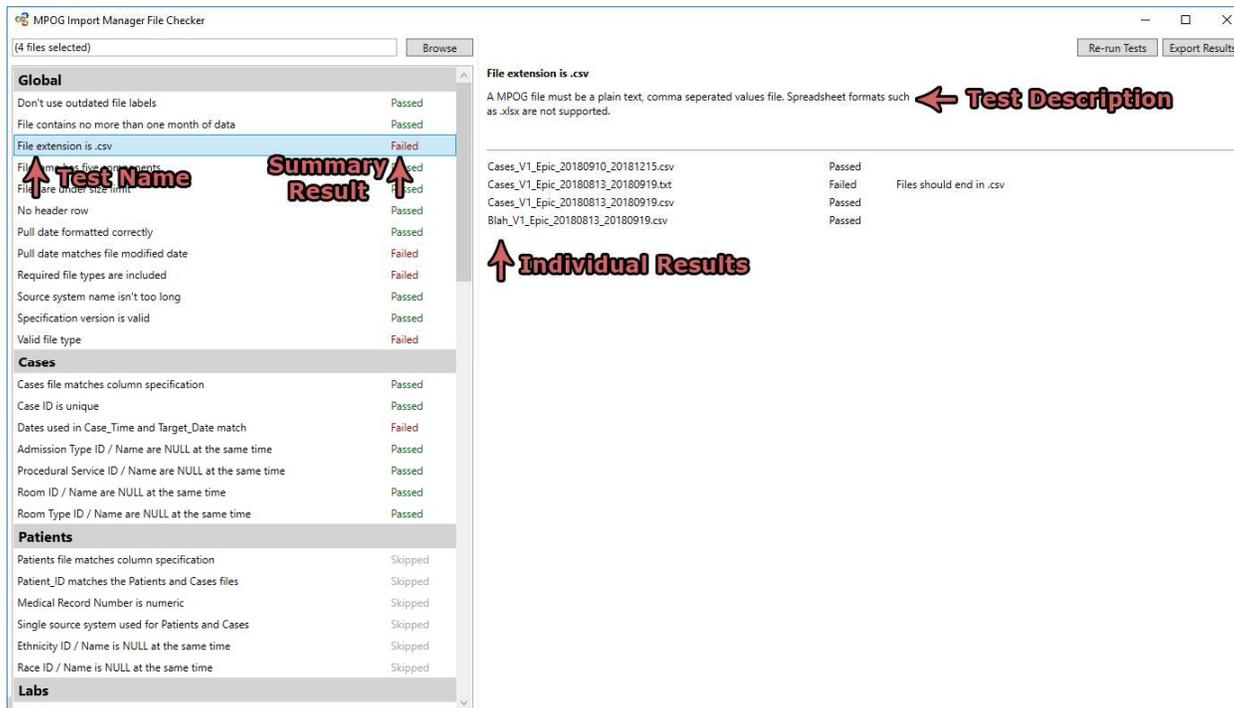
Due to memory limitations, do not generate files with more than one month's worth of data.

Furthermore, it is strongly recommended to begin testing using files that contain only one day's worth of data. Doing so will allow both the extracts and File Checker to run more quickly, leading to faster iteration.

Reading the Results

The application lists all tests it performs on the left side of the application window. Each test in the list belongs to a category. These categories reflect the different file types found in the specification. Additionally, there is a "Global" category which are tests that apply to all files.

Once the application has finished scanning files, each test will display a summary result which reflects the most severe issue the test encountered. Clicking on a test in the list will bring up more detailed information regarding that test's results. This information is displayed on the right side of the application.



Results in the File Checker interface

The test detail area contains the test name (bold text), a brief description of what the test checks for and a list of individual results. For each file that the test is applicable to, the individual results will contain at least one row. The same file can have additional rows if multiple errors or warnings were detected.

Warning: Not all violations will be listed

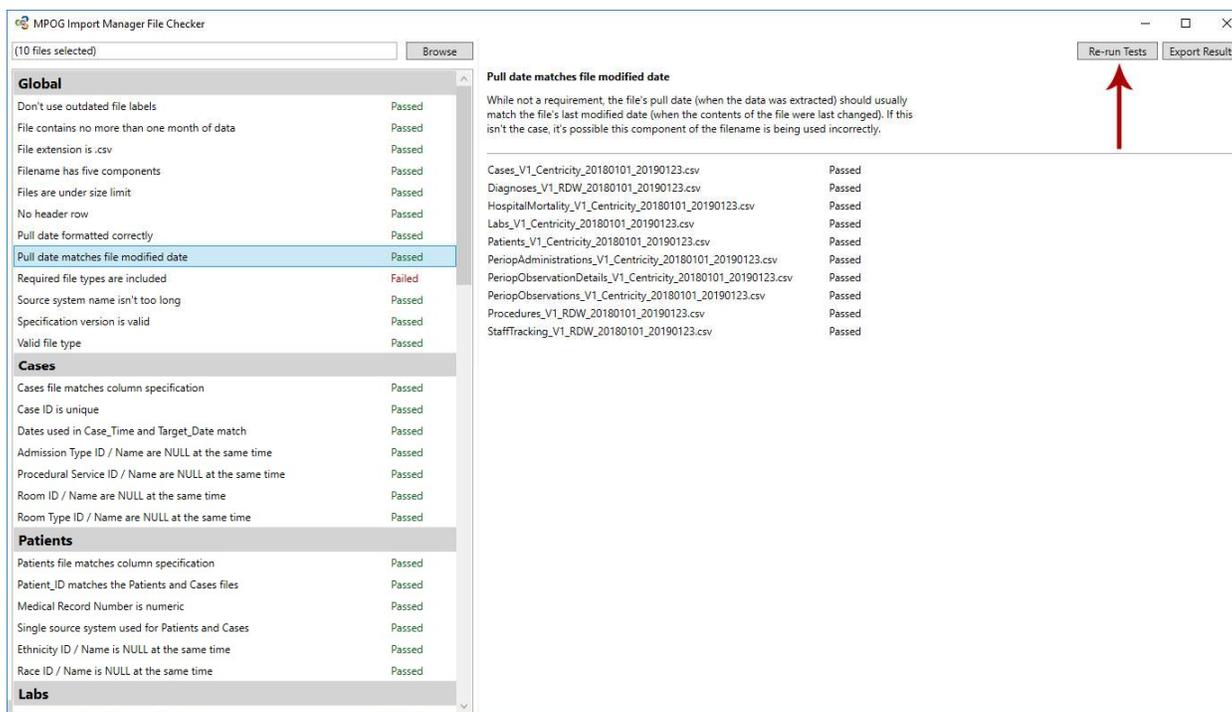
A test will only report only up to ten failures/warnings per file. Files may contain additional suspect rows even though they are not listed.

Tests can return a variety of result types, which are described in the table below.

Type	Description
Passed	The test found no issues in the file.
Warning	The test found suspect data in the file. This type of result does not necessarily indicate the file is incorrect, but the specified elements warrant investigation.
Failed	The test found a violation of the file specification or a data trend that is very likely incorrect.
Skipped	The test was not relevant to the set of files loaded.
Error	The data in given file caused the test to abort after encountering an unexpected error. This result likely indicates a bug in the File Checker logic.

Re-run tests

While testing a file extract you may overwrite an existing file with a new one with corrected data. If you already have the file you'd like to test selected in File Checker, you can either 1) re-load the files as described earlier or 2) click the "Re-run Tests" button in the top right corner of the application window. This will cause all tests to scan the previous selection of files again without having to use the "Browse" window.



Button to re-evaluate currently loaded files

Generating a result report

In order to easily share the results generated by File Checker, the application has the ability to export a summary report. Click on the “Export Results” button in the top right corner of the application window and save the report to the desired location. You can then share this report either internally or with your MPOG support contact.

The screenshot shows the MPOG Import Manager File Checker application window. The title bar reads "MPOG Import Manager File Checker". Below the title bar, there is a text field containing "(10 files selected)" and a "Browse" button. In the top right corner, there are two buttons: "Re-run Tests" and "Export Results". A red arrow points to the "Export Results" button.

The main content area is divided into two panes. The left pane shows a list of test results, grouped by category:

- Global**
 - Don't use outdated file labels: Passed
 - File contains no more than one month of data: Passed
 - File extension is .csv: Passed
 - Filename has five components: Passed
 - Files are under size limit: Passed
 - No header row: Passed
 - Pull date formatted correctly: Passed
 - Pull date matches file modified date: Passed**
 - Required file types are included: Failed
 - Source system name isn't too long: Passed
 - Specification version is valid: Passed
 - Valid file type: Passed
- Cases**
 - Cases file matches column specification: Passed
 - Case ID is unique: Passed
 - Dates used in Case_Time and Target_Date match: Passed
 - Admission Type ID / Name are NULL at the same time: Passed
 - Procedural Service ID / Name are NULL at the same time: Passed
 - Room ID / Name are NULL at the same time: Passed
 - Room Type ID / Name are NULL at the same time: Passed
- Patients**
 - Patients file matches column specification: Passed
 - Patient_ID matches the Patients and Cases files: Passed
 - Medical Record Number is numeric: Passed
 - Single source system used for Patients and Cases: Passed
 - Ethnicity ID / Name is NULL at the same time: Passed
 - Race ID / Name is NULL at the same time: Passed
- Labs**

The right pane shows a list of files and their status:

File Name	Status
Cases_V1_Centricity_20180101_20190123.csv	Passed
Diagnoses_V1_RDW_20180101_20190123.csv	Passed
HospitalMortality_V1_Centricity_20180101_20190123.csv	Passed
Labs_V1_Centricity_20180101_20190123.csv	Passed
Patients_V1_Centricity_20180101_20190123.csv	Passed
PeriopAdministrations_V1_Centricity_20180101_20190123.csv	Passed
PeriopObservationDetails_V1_Centricity_20180101_20190123.csv	Passed
PeriopObservations_V1_Centricity_20180101_20190123.csv	Passed
Procedures_V1_RDW_20180101_20190123.csv	Passed
Stafftracking_V1_RDW_20180101_20190123.csv	Passed

Below the file list, there is a text box with the following text: "Pull date matches file modified date. While not a requirement, the file's pull date (when the data was extracted) should usually match the file's last modified date (when the contents of the file were last changed). If this isn't the case, it's possible this component of the filename is being used incorrectly."

Export results to a report file

What to Do Next?

Now that you have feedback on your data files and how well they adhere to the MPOG specification, the next steps will depend on the results returned by File Checker.

All Tests Passed or Skipped

Once you have a full set of files that can pass all tests, you are ready to continue with the MPOG setup phase. Please export a result summary report and send it to your MPOG support contact. MPOG will then validate your results.

Take careful consideration of the “Required file types are included” test. As its name suggests, this indicates which file types are necessary to proceed. File types not listed here are optional and it is acceptable to have results of “Skipped” in these categories. Such file types are crosswalk and payer files.

Encountered Warnings

You may encounter warnings if the data scanned is suspect. This can happen when match rates are on a borderline threshold, or if the test is not able to confirm a violation of the specification. Any warnings should be investigated for data accuracy. Please document the cause for a given warning as well.

Encountered Failures

Failures indicates violations of the file specification or high-certainty data inaccuracies. Any failures will prevent the Import Manager framework from accepting the files and therefore must be resolved before proceeding to the MPOG setup phase.

Encountered Errors

If any of the tests return an “Error” result, please notify your MPOG support contact as it indicates a bug in the application.

The application will write to a log file (error.txt) in the same location as the executable. This will contain additional information regarding any errors. Please include this log file if it is available.

Application Appears to Be Stuck

The application may take a long time to scan a file if it contains a large amount of data. However, if you suspect the application is stuck look to the bottom left corner while the application is processing to see which file is causing the issue. The application should be able to evaluate a set of files well within five minutes, so reach out to your MPOG support contact if you are encountering poor performance.