



# MHDO Data Warehouse Submission Validation Process: Overview

April 17, 2013

## Summary

This document outlines the planned data submission validation process that will be implemented for the MHDO data warehouse. Its intent is to provide a high-level overview of the validation process. The conceptual details of the process outlined here represent the current plan for warehouse development; however the details of how this process will actually be implemented remain to be specified. The actual details of how this functionality is implemented may differ from the conceptual flow outlined here, although the same functionality will be provided.

Currently, elements of this process have been referred to as “edits” or “business rules.” This terminology will not be used in this document in favor of the terms “validation” and “validation logic.”

## Validation Process

The validation process can be thought of as a series of evaluations that are performed on data before it is allowed into the data warehouse. The process will perform as many of the validations as it can each time it is run to ensure that the list of issues that it generates is as comprehensive as possible. Once this list of issues has been generated, a resolution must be determined for each issue. This resolution might take the form of warning overrides, file resubmission, or threshold exemptions. Only when all issues have been resolved will a file be able to be accepted.

## Four Levels of Validations

There are four levels at which submitted data will be evaluated. These are:

- Structural validation
- Record-level content validation

- File-level content validation
- Submission-level content validation

### **Structural Validations**

Before any of the content of a file or record can be validated, we must first be able to parse and interpret the data. Structural validations will be done at the file-level and at the record-level.

#### **Examples**

File-level structural validations will look for issues such as header totals not matching the total number of records or issues that prevent the type of file from being determined.

Record-level structural validations will look for issues such as the number of fields not matching the expected number or field data having invalid formats.

### **Record-level Content Validation**

Record-level content validations will evaluate the data contained within an individual record and determine whether it adheres to validation logic that defines both the required and the expected characteristics of records. Only issues that can be detected by looking at a single record will be detected at this level. The two main types of validations will be code content validations and internal consistency validations.

#### **Examples**

Code content validations will check the contents of a given field against an appropriate validation set, such as valid CPT/ICD9 codes or ZIP codes. Internal consistency validations will check the contents of one field against the contents of one or more other fields.

### **File-level Content Validation**

File-level content validations will evaluate the contents of an entire file rather than just an individual record. The two main types of validations will be statistical property validations and cross-record consistency validations. Statistical properties can be defined as an average value for a field across a file, the percentage of qualifying cases passing a Boolean test, or other similar types of calculations.

### **Examples**

Statistical property validations will determine if a statistic, such as the percentage of cases that are associated with males, is within a pre-defined threshold. Cross-record consistency validations ensure that values that should be consistent within a file, such as the service date and payment date are, in fact, consistent.

### **Submission-level Content Validation**

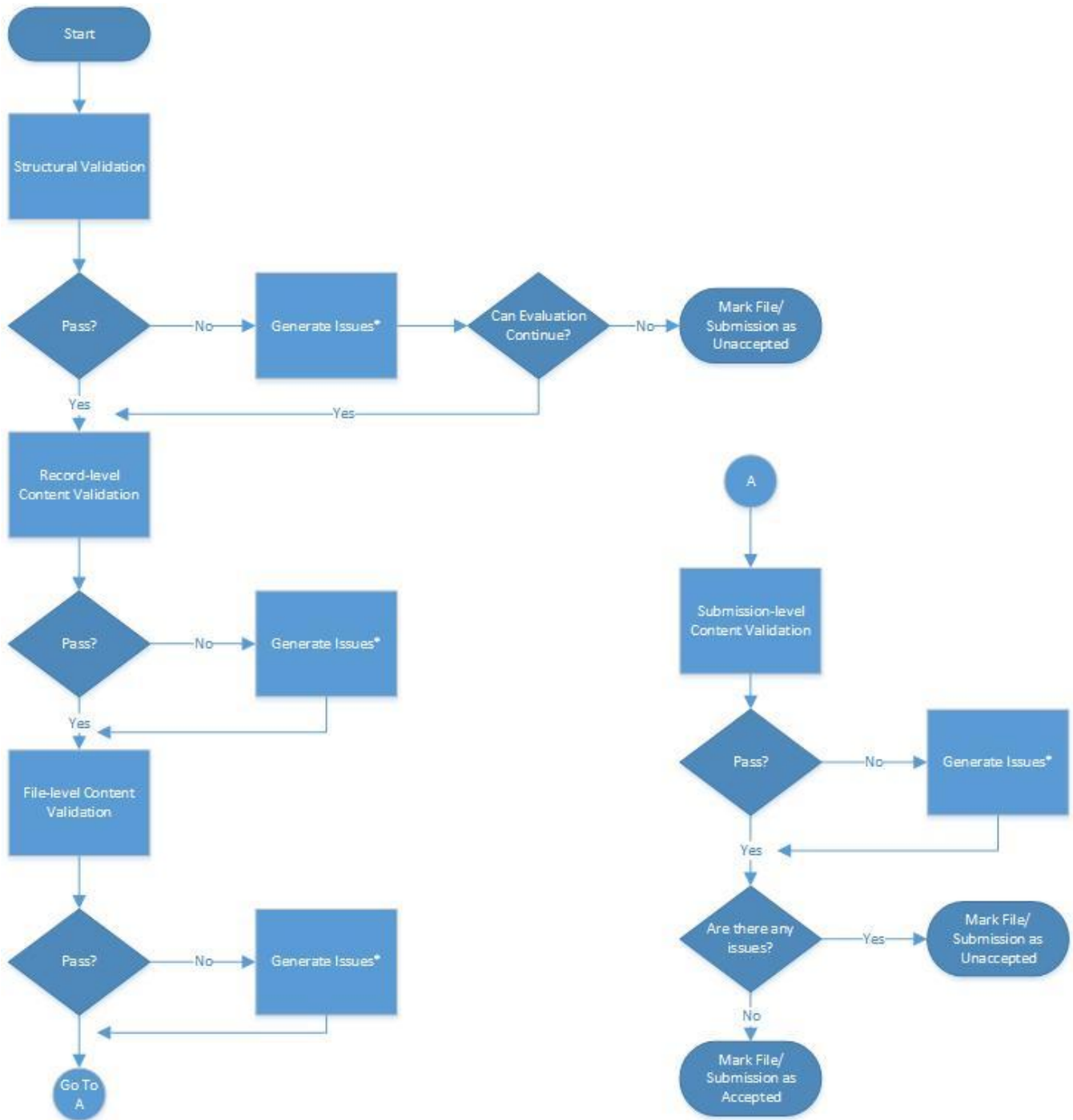
Submission-level content validations are similar to file-level validations, except that they consider data across all files submitted during the submission period, and in the future, across historical data submissions. Characteristics of data can be calculated and compared to thresholds to determine whether the validation is passed or not. Currently, there are two main types of submission level validations: statistical property and entity-matching validations.

Once historical claims data have been migrated to the warehouse, significant departures from historical norms will be flagged independent of the formal thresholds., e.g., if claim types fluctuate dramatically from month to month.

### **Examples**

At the submission level, statistical property validations will determine if, for instance, the average claim value per member is within a pre-defined threshold. Entity-matching validations will ensure, for instance, that every claim has a corresponding eligibility record submitted.

### Sample Validation Process Flowchart



\* Issues generated during each stage of the process will be added to validations report produced at the end of the validation process.

## Issue Resolution

When any of the validations outlined above fail, an issue will be generated. Each issue will contain full detail on the type of failure encountered, as well as detailed information on all the fields, records or files involved in the failure. In order for a file (or a set of files, when talking at the submission-level) to be accepted, there must be no unresolved issues.

The most basic way that an issue may be resolved is via file resubmission. When a data submitter determines that they can resolve an issue on their end, they may elect to do so and resubmit a file. When a file is resubmitted, the full set of validations is rerun. This may or may not result in a new set of issues to resolve (the old issues are superseded by the new ones). By comparing the two sets of issues on each file, MHDO will be able to evaluate whether the data submitter was able to successfully resolve any issues.

If issues are not fixed by file resubmission, the options available for resolution depend on the type of issue encountered.

## Four Types of Issues

Each validation will be associated with an **issue type**. This issue type will determine the options a data submitter has when a validation fails to pass. The issue type also determines whether the thresholds associated with a given validation may be overridden by values in the data submitter's profile or not.

The four types of issues are:

- Failure-level
- Exemption-level
- Ad hoc-level
- Profile-level

## Minimum Passing Thresholds

Each validation will have default success criteria defined which specifies what must be true in order for the validation to be passed. In some cases, this will be a simple true/false determination, such as "can we open and process the file?" However, in many other situations, these criteria will take the form of percentage thresholds of records within a file which must pass the evaluations.

The percentage threshold associated with a given validation is called the minimum passing threshold for the validation. It is calculated as follows:  $100 * (\text{count of records passing the validation}) / (\text{count of records qualified for the validation})$ .

### **The Data Submitter Profile**

Each data submitter will have a “profile” in which these default criteria can be changed for their submissions. Depending on the type of issue (see below), changes to the data submitter profile may either be made directly by the data submitter or by MHDO via an exemption request. In either case, though, MHDO will have the ability to change or reset any success criteria in a data submitter profile based on its policies and procedures.

### **Failure-level Issues**

Failure-level issues are the most serious type of issue. These are issues that must be resolved in order for a file to be accepted. They will typically prevent the use or evaluation of the data provided. The only way these types of issues can be resolved is through resubmission of the file that has the issue.

### **Exemption-level Issues**

Exemption-level Issues will generally be for fields required for “Prompt Payment of Clean Claims.” MHDO may agree to bypass certain validations or negotiate thresholds different than the default minimum passing threshold on a submitter-by-submitter basis. These “exemptions” from the default thresholds will typically be for a limited period of time; however, exemptions can be made permanently at MHDO’s discretion.

If a file or submission generates an exemption-level issue, the submitter may apply for an exemption (either for that one time or for a limited period) via the MHDO portal. MHDO will then review the exemption request and, if approved, clear the outstanding issue.

Exemptions that allow a submission to pass a validation it would have otherwise failed will be documented in the metadata that is associated with the data in the data warehouse.

### **Ad hoc-level Issues**

Ad hoc-level issues are issues that submitters may override, but that MHDO does not allow to be set permanently in the submitter’s profile. These validations will always use the default thresholds

defined by MHDO. Data submitters will be able to override these warnings on a case-by-case basis via the MHDO portal. Submitters will be required to provide a reason for the override.

Overrides to issues made by data submitters will be documented in the metadata that is associated with the data in the data warehouse.

### **Profile-level Issues**

Profile-level issues represent validations that vary by book of business. Data submitters do not need to apply for an exemption to change the minimum passing threshold in their data submitter profile, they will be able to make these changes themselves via the MHDO portal, along with a reason for the change. Once a profile change is made, the submitter's data will no longer be subject to the specific validation. However, MHDO will have the ability to change or reset data submitter profile thresholds that it deems to be inappropriate given its policies and procedures.

Data submitter profile thresholds different from the defaults that allow a submission to pass a validation it would have otherwise failed will be documented in the metadata that is associated with the data in the data warehouse.