HEDIS Project Specification

Massachusetts Health Data Consortium

Version 1.0.0, 2019-12-31

Table of Contents

The first project of the Massachusetts Health Data Consortium (MHDC) Data Governance Collaborative is determining a common format for sending HEDIS data between payers and providers. Version 1.0.0 of the HEDIS data exchange involves transfer of text files via secure FTP supported by mutually agreed upon data and file formats.

It includes documentation describing the general process and guidelines for HEDIS data exchange under the main heading Process. It includes documentation describing the file transfer process and the file format under the main heading File Exchange. It includes documentation describing the organization of data and the rules for each field under the main heading Data Dictionary.

Process

This section includes information related to the general process of exchanging HEDIS data.

Process Overview

The Massachusetts Health Data Consortium Data Governance Collaborative HEDIS project version 1.0.0 is designed as the first step in the process of standardizing and automating data exchange between payers and providers. In this version, we are focused on beginning the standardization process. The actual data exchange will continue to be manual using flat text files exchanged over secure FTP (providers may automate the secure FTP process in some way if they so desire, but such automation is outside the scope of this specification).

This version of the project does not need to be 100% standardized to meet its goals. The idea is to start down a path that will eventually lead to standardized data and automation. As such, in cases where full agreement was difficult or the system could easily be set up to accommodate some flexibility this path was chosen to expedite taking the first steps in the process more quickly.

An example of this is the use of valid values file to indicate exactly what data is expected by a payer from each provider. Instead of defining a long list of enumerated values for certain fields, the allowed values are determined by these files and are manually enforced by people at both ends of the transfer. Thus, the high level HEDIS process is as follows:

1. A payer puts a valid values file and a list of desired patients on the secure FTP server

2. The requisite provider downloads these files

3. That provider gathers the requested data on those patients

4. That data is formatted into the format described by the HEDIS standard

5. The resulting files are spot checked per the rules defined in the project standard

6. Files are fixed as needed and checked again

7. Once files pass they are posted to the secure FTP server

8. The payer downloads those files and uses the information to calculate HEDIS scores

> **ℹ** Various steps of this process will be discussed in more detail in later sections of the documentation

Basic Rules

The following basic rules guide the data collection process:

- data is collected over the course of the current calendar year and for previous calendar years as indicated in the data range rules outlined in this specification

- data cutoff for each file is the end of a calendar month

- data request files do not roll over; new files must be sent for each data request even if the only change is the deadline

- providers must have at least one week to process data before it's due

- providers must have at least ten days after the end of a month to process data for that month

- data is cumulative over the course of the year; new data collected is appended to the data sent in previous months prior to sending the entire data set to payers (if it is easier for a provider they may choose to collect the entire data set from scratch each time rather than physically appending new records into existing files so long as the result is the same; i.e. that each new set of data contains all of the available data from the entire relevant time span)

- data for patients who are not still active in the plan at the time of the request may or may not be included; providers will include any requested data they have available but may not retain access to all prior information or be able to use old member IDs to access all historical information

- any data that does not fit into the prescribed rules and guidelines of this specification should be immediately brought to the attention of MHDC so a solution can be found and added to the specification.

Data Range Rules

Although HEDIS evaluation is an annual process, the data used for the relevant calculations predates the current year. The following data must be sent for full compliance with the MHDC HEDIS specification:

- All patient data for the last three completed calendar years

- All patient data for the current calendar year through the last completed month that is 10 or more days before the requested data exchange date. Data sent on October 8 should include data through the end of the month of August while data sent on October 11 should include data through the end of the month of September

- All applicable cervical cancer screening tests performed within the past 5 years

- All applicable colorectal cancer screening tests performed within the past 10 years

Applicable Cervical Cancer Screening Tests

The following tests are considered applicable cervical cancer screening tests:

- pap smear/cervical cytology

- human papillomavirus (HPV)

Applicable Colorectal Cancer Screening Tests

The following tests are considered applicable colorectal cancer screening tests:

- colonoscopy

- flexible sigmoidoscopy

- fecal occult blood test

- computed tomography colonography

- stool DNA test

Data for 2020

Data sent in 2020 would include:

- patient data for 2017, 2018, and 2019
- patient data for 2020 as the year progresses and it becomes available
- All pap smears performed since 2015, if any
- All human papillomavirus tests performed since 2015, if any
- All colonoscopies performed since 2010, if any
- All flexible sigmoidoscopies performed since 2010 , if any
- All fecal occult blood tests performed since 2010, if any
- All computed tomography colonographies performed since 2010, if any
- All stool DNA tests performed since 2010, if any

Case Considerations

This documentation will address casing rules throughout as different elements of data, files, and fields are discussed. In general, preference is given to using camelCase for most things, meaning that multiple word names or string values are pushed together into a single segment where the initial letter is lower cased but the start of all other portions of the new compound word is upper cased. For example, the following are all using camelCase:

- firstName
- pacificIslander
- isEnabled
- patientId

Case Sensitivity

In general, all data and names are assumed to be case sensitive. However, values that differ only by case are not permitted. Any exceptions to this will be noted as relevant throughout the documentation

When Case Has Meaning

If case has meaning, it may be retained even if the rules for a specific piece of data would indicate otherwise more generally so long as it does not violate the casing of a supplied enumerated value or unit designator (enumerated values and unit designators must always match the documented values exactly including case). For example, proper nouns may have their first letter capitalized even if general guidelines indicate that letter should be lower cased. Conversely, names of fields will always be presented in camelCase in the documentation even in headings or if starting a sentence (but sentences will generally be written to avoid this situation).

Per industry standards, acronyms or abbreviations which are typically presented in all capital letters are not covered by this exception; they should use the case as indicated by the specification. Note in the examples of camelCase above, the abbreviation ID is specified as *Id when it is the second term of a compound camelCase name. Similarly, a field named id is correct nomenclature for a field containing an ID.

Patients

Each patient covered by any insurance plan must have a patient ID number based on their member ID and personal information. See Representing Patients with Patient IDs for specific information on patient IDs.

> If MHDC changes the ID format in the future, it may be necessary for payers to generate lists of patient IDs and transmit them to providers. In that case, new patients files may be needed throughout the year as addendum covering patients who joined a payer plan mid-year.

Identifying Patients to Include

There are two options supported:

1. Send data for all patients on any payer plan
2. Send data for only the patients payers request

Each payer may decide which option it wants each provider to use. A payer is not required to use the same option for each of its providers.

Regardless of the option chosen, each data exchange contains data through the end of a particular calendar month.

Send Data for All Patients

If a payer wants a provider to send data for all patients using one of their plans, no action is needed on a monthly basis. The valid values file sent by the payer to the provider will define the deadlines for data exchange.

Send Data for Specific Patients

If a payer wants a provider to send data for specific patients using one of their plans, a file listing the requested patients should be uploaded to the secure FTP server.

This file should conform to the file naming conventions outlined here

if present, list the following in order in pipe-delimited rows:

- patient ID
- subscriber number
- first name

- last name
- date of birth in `YYYY-MM-DD` format

The date included in the file name is a deadline for the related return data. This date must be at least one week after the patients file is uploaded and ten days after the end of the calendar month. It must agree with the date of the corresponding valid value files or both files should be rejected.

Files with deadlines that do not provide a seven day window to respond will be rejected. Files with deadlines that meet the first requirement but not the second should result in sending data for the previous calendar month. For example, a patients file uploaded on September 30 with a deadline of October 8 should be fulfilled with data through the end of August. However, a patients file uploaded on September 30 with a deadline of October 10 should be fulfilled with data through the end of September.

> ℹ️ Version 1.0.0 of the MHDC HEDIS specification does not provide a common notification mechanism and allows the option of scheduling actions on a regular basis (i.e. for the same time each month). If a payer-provider pair is using this option, the payer must still conform to the date requirements in file names and the times for action outlined in this documentation. See Notifications for more information.

Patients Who Leave a Plan Mid-Year

Once a patient is no longer part of a plan some or all of their data may no longer be associated with their former member ID and thus not accessible when using that ID to determine the data to include in a file transfer.

Each provider will make an effort to include as much of the relevant data from the time period when a patient was covered by a particular payer when preparing data to transmit to that payer. However, this data may be incomplete or even entirely unavailable.

Patients Who Join a Plan Mid-Year

When new patients join a plan in the middle of the reporting year their data should be included (if requested or if all patients are requested) for the covered dates in any files generated after they join the relevant plan.

Valid Values

The Massachusetts Health Data Consortium HEDIS specification allows for some flexibility in the data exchanged between payers and providers. This flexibility falls into two categories:

- flexibility in which fields in the data dictionary are required to contain data (data may be sent in other fields as optional additions that may be ignored by the payer)
- flexibility in the specific diagnoses, laboratory services, medications, observations and vital signs, immunizations, procedures, and exceptions to be sent for patients. This flexibility is supported independently for the services outlined in each data group except for memberships.

This flexibility is handled via a valid values file provided by each payer to each provider outlining the requirements in the categories above. This file is part of the data exchange contract; items listed in the file must be provided for any patient with extant data of that sort. For example, not every patient included in a data set will have A1c scores but if A1c is listed in a valid values file then A1c scores must be provided for every patient in the data set who has them. The MHDC HEDIS specification does not permit payers to specify which data to send for which patients.

> This file does not change the available fields as defined in the MHDC HEDIS specification in any way. It merely indicates which fields a particular payer wants a particular provider to populate and the services it supports for each relevant data group.
>
> If a valid values file includes fields or services that are not part of or supported by the specification ignore them (do not include the field in any way) and make sure the payer is aware of this action so they properly align the data to the fields that are defined. If the payer determines these fields or services are important, they should immediately contact MHDC so they (or any relevant units or other necessary items) can be added to the specification.

This file should conform to the file naming conventions outlined here

Deadlines in Valid Values Files

The date included in the file name of a valid values file is a deadline for the related return data. This date must be at least one week after the valid values file is uploaded and ten days after the end of the calendar month. If a corresponding patients file is also supplied the dates of both files must agree.

Files with deadlines that do not provide a seven day window to respond will be rejected. Files with deadlines that meet the first requirement but not the second should result in sending data for the previous calendar month. For example, a valid values file uploaded on September 30 with a deadline of October 8 should be fulfilled with data through the end of August. However, a valid values file uploaded on September 30 with a deadline of October 10 should be fulfilled with data through the end of September.

> Version 1.0.0 of the MHDC HEDIS specification does not provide a common notification mechanism and allows the option of scheduling actions on a regular basis (i.e. for the same time each month). If a payer-provider pair is using this option, the payer must still conform to the date requirements in file names and the times for action outlined in this documentation. See Notifications for more information.

Contents of Valid Values Files

Include the following in a valid values file:

- lists of expected services in each data group
- lists of expected fields in each data group file

These lists should be underneath the name of the type of content or the pertinent data group (membership, observations, etc) as indicated by the list of headings below and be included as follows in the syntax section below, one data point per line.

Categories in Valid Values Files

Each list above is considered a single category of the valid values file and contains data for either the services to include within each data group or the fields that should be populated in each data group. These two types of categories have very different behavior.

If a particular category listing services to include for a data group is not included in the valid values file then all data for all supported services in the category must be sent. If nothing is required for a particular category then omit the category but also omit the related data group (see below).

Conversely, if a particular category listing fields to include for a data group is not included in the valid values file then no fields are expected for that data group. In this case, the related file should be included in the data transfer as an empty file.

> The inclusion of a services list for a particular data group in the valid values file does not eliminate the need to separately specify which fields to include in the data groups themselves; they are orthogonal requirements. However, if a list of fields is not included for a particular data group then any list of services for the same data group is moot and should be ignored.

Rows containing data (either fields or services) not listed for inclusion in the data set per the valid values file will be ignored by the payer but will not cause errors or trigger fault tolerance levels when evaluated by the payer.

The patient ID must be included in the field list for every data group that includes other fields.

A payer can decide what to include for each category independently and can choose different options for different providers. The included options may change on a month by month basis.

Valid Value File Format

A valid value file consists of a series of groupings containing a heading followed by the list of fields or services pertinent to that heading. Each of these groupings represents a single category as described above.

Each grouping must be separated by at least one empty line (additional empty lines should be ignored) and no empty lines should appear within any group's data. The lists of expected services should be first followed by the lists of required fields for each data group (the data groups may appear in any order in both lists as long as all of the services appear before any of the field categories do). The list of fields within each data group should be presented in the order they appear within the data group in a row of data; if the first, second, and fifth fields are required they should be listed in that order.

Headings

The headings for the list of expected services for each data group are an abbreviated term for the

type of data followed by a colon as indicated by the table below. The headings for the list of required fields for each data group are the data group names with the first letter capitalized followed by a colon.

Table 1. Valid Value Services Category Headings

Heading	Data Group	Notes
Diags:	diagnoses	A list of diagnoses that should be included in the data exchange when they apply to any patients in the data set within the normal data range. If the category is not included in the valid values file all known diagnoses should be sent for each patient.
Exs:	exceptions	A list of exceptions that should be included in the data exchange. If the category is not included in the valid values file all known exceptions should be sent.
Labs:	laboratoryServices	A list of laboratory tests that should be included in the data exchange when they have been performed within the normal data range on any patients in the data set. If the category is not included in the valid values file all known laboratory results for all supported laboratory services should be sent.
Meds:	medications	A list of medications that should be included in the data exchange when they have been prescribed within the normal data range for any patients in the data set. If the category is not included in the valid values file data for all known prescriptions should be sent.
Procs:	procedures	A list of procedures that should be included in the data exchange when they have been performed within the normal data range on any patients in the data set. If the category is not included in the valid values file data for all known procedures should be sent.

Heading	Data Group	Notes
Vax:	immunizations	A list of immunizations that should be included in the data exchange when they have been performed within the normal data range on any patients in the data set. If the category is not included in the valid values file data for all known immunizations should be sent.
Vitals:	observations	A list of vital signs that should be included in the data exchange when they have been performed within the normal data range on any patients in the data set. If the category is not included in the valid values file all data for all supported vital signs should be sent.

Rows

The lists of expected services and the list of expected fields for a data group use a pipe ("|") delimiter to separate individual fields within each row. No delimiter is used to indicate the row start or row end positions; if a row ends with a delimiter it indicates that the last field in the row is blank.

However, the two sets of lists convey different information and thus look very different. Each list of expected services has three values per row while each list of fields has only one. Thus, although technically a pipe-delimited row of data, no delimiters actually appear in a list of field data and it amounts to a simple text list with one item per row.

Each row of data in a list of expected services contains the following in the order listed below:

1. A description or name providing some context for the services
2. The code system expected for the service
3. The allowed code for the service

If a particular service may use more than one code or more than one code system, each code system-code pair should be given a separate line in the list.

If the payer does not care which code system is used as long as it's one supported by the MHDC HEDIS specification then the code system and allowed code fields should be blank.

> This was an executive decision made by MHDC during the final update as the question had not been addressed in the related discussion during the review meeting. If you do not want to support this option, do not use it in any of your valid values file.

Example Valid Values File

The following valid values file might be used in a payer request for data:

```
Vitals:
height|loinc|8302-2
height|loinc|3137-7
height|snomed|50373000
weight|loinc|29463-7
weight|snomed|27113001
systolic < 130|cpt2|3074F
systolic 130-139|cpt2|3075F
systolic >= 140|cpt2|3077F
systolic|loinc|8480-6
systolic|snomed|271649006
diastolic < 80|cpt2|3078F
diastolic 80-89|cpt2|3079F
diastolic >= 90|cpt2|3080F
diastolic|loinc|8462-4
diastolic|snomed|271650006
temperature|snomed|386725007
temperature|snomed|431807005
pulse||

Labs:
a1c|cpt|83036
a1c|mapped|a1c
a1c|snomed|365845005
potassium|cpt|001180
potassium|loinc|6298-4

Vax:
rubella and mumps|cvx|38
measles, mumps, rubella|snomed|38598009

Memberships:
patientId
lastName
firstName
gender
birthDate

LaboratoryServices:
patientId
serviceCode
codeSystem
serviceDate
result
unit

Observations:
patientId
serviceDate
observationCode
```

```
codeSystem
value
unit
```

In this case, blank files would be sent for all data groups except memberships, laboratoryServices, and observations including for immunizations. That's because while services are defined for immunizations, no fields are defined for the immunizations data group. Therefore, no immunization fields should be populated and sent to the payer by this provider.

File Exchange

HEDIS information is sent by each provider to each payer monthly by secure FTP in the format described in this documentation.

File Structure

The HEDIS data is currently split into groups of related content sent in separate flat text files containing rows of data each representing a specific instance of that type of data associated with a specific patient.

These files use a pipe ("|") delimiter to separate individual fields within the row. No delimiter is used to indicate the row start or row end positions; if a row ends with a delimiter it indicates that the last field in the row is blank.

The fields are order dependent, meaning they must be presented in the order listed in the data specification.

In addition, there are several helper files that provide information about patients and data requests to the provider.

File Names

The HEDIS data exchange has several different types of files. Files sent from the payer to the provider may include:

- Patients Files
- Valid Values Files

These files are control files that indicate the patient data to send.

Files sent from the provider to the payer contain the actual patient data in Data Group Files

Files that do not conform to this naming convention will be rejected by the sender and the recipient will be expected to provide new versions that do conform within three days.

Using Directories

The Massachusetts Health Data Consortium HEDIS standard assumes that each payer maintains a separate directory on its secure FTP server for each provider. All files from that provider are identified as being for or from that provider by virtue of their placement inside this directory and thus the file names do not need to include the provider name for clarity. More information about the directory structure can be found here.

Patients Files

These files are used to request specific patient data from a provider.

These files should be named `patients_for`*Payer*`_date`

where *date* is the deadline for receiving the data in `YYYY-MM-DD` format and *Payer* indicates the payer requesting the data as follows:

- BCBS - Blue Cross
- Tufts - Tufts Health Plan
- Fallon - Fallon
- HPHC - Harvard Pilgrim

for example, patients_forTufts_2019-10-15

More information about patients files is available in the Send Data for Specific Patients section of the documentation.

Valid Values Files

These files are used to indicate what data a payer expects to receive from a provider.

These files should be named `validValues_for`*Payer*`_date`

where *date* is the deadline for receiving the data in `YYYY-MM-DD` format and *Payer* indicates the payer requesting the data as follows:

- BCBS - Blue Cross
- Tufts - Tufts Health Plan
- Fallon - Fallon
- HPHC - Harvard Pilgrim

for example, validValues_forHPHC_2019-11-15

More information about valid value files is available in the Valid Values section of the documentation.

Data Group Files

The files containing the actual HEDIS data are split into data groups representing different types of data. The current data groups are:

- diagnoses
- exceptions
- immunizations
- laboratoryServices
- medications
- memberships
- observations

- procedures

These files should be named `dataGroup_forPayer_date`

where *dataGroup* is the camelCase name of a valid data group (listed above), *date* is the deadline for receiving the data in `YYYY-MM-DD` format, and *Payer* indicates the payer requesting the data as follows:

- BCBS - Blue Cross

- Tufts - Tufts Health Plan

- Fallon - Fallon

- HPHC - Harvard Pilgrim

for example, laboratoryServices_forHPHC_2019-11-15

More information about the data groups and the data expected in each group's file is available in the Data Dictionary specification.

Directory Structure

Rather than just dumping a slew of flat files into a single area for processing which can be time consuming to scroll through and onerous to maintain, the MHDC HEDIS data exchange specification is adopting a directory structure for its adopters to use.

Our basic approach to defining a directory structure is:

- To clearly identify space specific on each payer's secure FTP server

- To provide a mechanism for archiving files that have already been addressed

- To provide a mechanism for clearly identifying rejected data

> In version 1.0.0 of the MHDC HEDIS specification, validation rules are recommended guidelines. Participating organizations are not required to strictly enforce validation or mark rejected data in any way. Thus, while recommended, there is no requirement that the mechanism for identifying rejected data be used.
>
> Also, while future versions may institute an errors directory or more strict marking of errors in files, currently this is beyond the scope of the specification.

At a high level, the directory structure will be used to do the following:

- Provide a clearly marked space for each provider using the provider name or a recognizable abbreviation.
  - All files for or from a provider are placed in this directory
  - Only files for or from the provider are placed in this directory
- Provide an archived directory under the main provider directory
  - Files should be archived for at least 12 months once consumed (longer is acceptable)

- Files for the provider should be archived by the payer after the provider fulfills the related requests to acknowledge receipt of the uploaded files
- Files for the payer should be archived by the payer once they have been successfully consumed
  - All rejected rows should be addressed before this happens
- Include a rejected directory under the main provider directory
  - Any files that are rejected because of the error rules or fault tolerance limits should be moved here
  - Files should remain here until substitute files are accepted. At that point they may be deleted or moved to an archived subdirectory of the rejected directory at the discretion of the payer
  - Files not rejected wholesale but with rejected rows should remain in the top level directory; rejected rows within those files should be marked as rejected by prepending a field containing R to the row

Basic Directory Structure

Each payer's secure FTP server should include a top level directory for each provider. This directory should have the same name across all payers and be commonly recognizable as representing the provider. The list of these directory names is maintained by MHDC and considered part of this specification. Generally, it will be a common short form of the provider's name with all spaces removed and all letters in lower case.

Table 2. Top-Level Directory Names

Organization	Short Forms	Directory Name
Berkshire Health System	Berkshire, BHS	berkshire
Beth Israel Lahey	Beth Israel, BIDMC, BIDCO, BIL, BILH, BILHPN	bethisrael
Boston Medical Center	BMC	bmc
New England Quality Care Alliance	NEQCA	neqca
Partners Health Care	Partners	partners
Steward Health Care	Steward	steward

Files from the payer to the provider

The payer currently provides either one or two files to the provider each time data is requested: valid values file and a list of desired patients. The valid values file is required and the list of desired patients is optional (if not provided, all patients will be included in the response data).

These files should be placed at the top level of the directory and should be archived per archival rules once the relevant data has been sent back. Currently there are no standardized rules for file upload or archival notifications. See Notifications for more information.

See Valid Values Files and Patients Files for more information about these files.

Files from the provider to the payer

The provider provides individual files containing the requested data in each data group.

These files should be placed at the top level of the directory and should be archived per archival rules once the files have been processed and accepted by the payer. Currently there are no standardized rules for file upload or archival notifications. See Notifications for more information.

See the Data Group Files documentation for more information about these files.

Resending Rejected Data

Any rejected files should be resent wholesale and _2, _3, etc should be appended to the file name to indicate version 2, version 3, etc until a version is fully accepted.

Files with rejected rows should be resent with just the corrected rows and _p2, _p3 should be appended to indicate partial version 2, partial version 3, etc - this indicates all archived files must be collated together to get a complete picture of the data should it be needed at a later date

Currently there are no standardized rules for file upload notifications for replacement files. See Notifications for more information.

> In version 1.0.0 of the MHDC HEDIS specification, validation rules are recommended guidelines. Participating organizations are not required to strictly enforce validation or mark rejected data in any way. Thus, while recommended, there is no requirement that this process for uploading corrections be used.

Example

The Partners area within Tufts Health Plan's secure FTP server be would a directory called partners that contains the following files after Tufts requests new data from Partners but before they respond:

```
archived
rejected
patients_forTufts_2019-10-15
validValues_forTufts_2019-10-15
```

Once Partners uploads the data requested for 10/15 but before Tufts consumes it in any way the directory might look like this:

```
archived
rejected
diagnosis_forTufts_2019-10-15
exceptions_forTufts_2019-10-15
immunizations_forTufts_2019-10-15
laboratory_forTufts_2019-10-15
medications_forTufts_2019-10-15
membership_forTufts_2019-10-15
observations_forTufts_2019-10-15
other_forTufts_2019-10-15
patients_forTufts_2019-10-15
procedure_forTufts_2019-10-15
validValues_forTufts_2019-10-15
```

And after Tufts acknowledges receipt of the files it might look like this (with the patients and valid values files moved under archived):

```
archived
rejected
diagnosis_forTufts_2019-10-15
exceptions_forTufts_2019-10-15
immunizations_forTufts_2019-10-15
laboratory_forTufts_2019-10-15
medications_forTufts_2019-10-15
membership_forTufts_2019-10-15
observations_forTufts_2019-10-15
other_forTufts_2019-10-15
procedure_forTufts_2019-10-15
```

And once all files have been processed and either accepted or rejected it might look like this:

```
- archived
- rejected
```

with all of the files either under archived or rejected.

However, there might be files with rejected data that were not rejected wholesale. If medications was rejected wholesale and laboratory had some rejected rows, the directory might look like this before any corrections are made:

```
- archived
- rejected
- laboratory_forTufts_2019-10-15
```

And the directory with replacement data present might look like this:

```
- archived
- rejected
- laboratory_forTufts_2019-10-15
- laboratory_forTufts_2019-10-15_p2
- medications_forTufts_2019-10-15_2
```

Note that the original laboratory file is still present because it was neither rejected wholesale or accepted. When the updated file is accepted, they should be moved to the archived directory together.

Notifications

In general, notifications should be sent to payers and providers when new files are uploaded they need to address and whenever file rejections or errors that need to be fixed occur. However, in version 1.0.0 of the MHDC HEDIS specification, the exact form of notification is left to each payer-provider pair to determine.

In addition, in lieu of notifications, scheduled timeframes for specific actions are allowed in version 1.0.0 of the specification provided both parties involved agree. If monthly schedules are used, file name conventions and requirements related to the amount of time each party has to perform specific actions must still be followed. For example, a regular data upload deadline on the 6th of the month means that uploads will not contain data from the just completed month but rather will be a month in arrears (per the requirement that providers have 10 days after the end of a month to process data from that month).

File Validation

In an ideal world, every field in every row of every file would be checked for accuracy in both format and content and the data sent would contain no errors. Unfortunately, that's not realistic in the real world in a primarily manual system.

While the ultimate validation responsibility rests with the payers, each provider is expected to spot check the files per the spot checking guidelines below for data that looks wrong in some way. This data may be in the incorrect format or may be in the correct format but just isn't sensible. Then, once received by the payer, the data is again checked per the error handling guidelines below (primarily for formatting issues) and either rows or entire files may be rejected based on the defined rules.

In version 1.0.0 of the MHDC HEDIS specification, these validation rules are recommended guidelines. However, participating organizations are not required to strictly enforce either the spot checking or error handling rules outlined below. They are provided as goals to strive toward and lists of potential problems to look out for throughout the data preparation and data absorption processes.

As the data exchange process matures, more validation will be required. Once the data exchange process involves automation much stricter validation will be essential. The plan is to require more validation as time goes on to make the transition to automation easier and not go from no validation to full validation in a single step.

Spot Checking Data

Manually checking the data within the current flat file system is an onerous, time consuming, and unpleasant process. However, there is no good way to automate this process at this time so some form of manual checking is needed as a sanity check and to determine how the process is working. Thus, the HEDIS standard is defining some rules for spot checking data in files before those files are transferred from the provider to the payer.

Goals

The following goals were used to guide the spot checking rules:

1. Catch recurring issues with data so they can be fixed both in the current files and in the preparation of future data files

2. Fix one-off errors in the data noticed while looking for recurring issues. Identifying and fixing every one-off error is not within the scope of this plan

3. Find the right balance between clean data and the monotonous manual checking of every bit of data so the process works fairly well but has a reasonable return on (time) investment

4. Reward providers for producing clean data by reducing the burden of spot checks over time when no issues are found

Approach

To expedite an inherently slow process, spot checks will not manually compare the data in the files to the source of the data. The assumption is that most errors will manifest themselves in other ways.

The biggest risk to this approach is that the ID is incorrect and thus the data is ascribed to the wrong patient. If deemed an unacceptable risk, this risk could be mitigated by checking just the ID against the actual data source. This will increase the time required to complete spot checks by an unknown but likely not trivial amount.

> We should not be overly concerned with the case where an incorrect patient ID is not assigned to an actual patient. This case should be detectable by the payer and cause a rejection of the row. Further, the consequences of associating the wrong data with a patient are more severe than not having a patient to associate with the data.

Data to Check

The following should be checked in each row identified for a spot check:

- correct delimiter - fields are separated by pipes.

- correct number of fields - there are the expected number of delimiters in each row

- order of fields - the order of fields in a row has meaning and thus must be preserved. If data is not in the expected order, it will be misinterpreted upon receipt.

- ID - the ID is in the correct format

- data type of fields - the data should be in the expected data type for the field.

  - Strings should not be inside quotations

  - Dates should be in YYYY-MM-DD format

  - Booleans should be either true or false

  - Numbers should use Arabic numerals with no leading zeros (except for values between 0 and 1)

- defined constraints of fields - the data should be within maximum value and character ranges, should be present if required, should be one of the valid enumerated values for enumerated fields, etc.

  > Case and spacing matters when checking enumerated values.

- fields with no data should be empty (the two pipes denoting the left and right boundary of the field data should be next to each other with no space between them)

- The row does not have a hard wrap, line breaks, or other control characters that would be interpreted as an end of line. It should extend out on a single line until completed.

- The data provided in a row falls within the set of data requested in any applicable valid values file.

- The data provided in a field matches the expected data for that field. If expected to be a valid blood pressure, it falls within the range of values possible for a blood pressure reading.

- The data provided in a field matches the expected units for that value. If the height is supposed to be in inches it does not appear to be in centimeters or in feet and inches.

- The unit designation provided in fields specifying units exactly matches one of the units supported by the MHDC HEDIS specification

> Unit designations are expected in exactly the format listed in the specification: %, per, Per, and percent are considered four different values for purposes of matching unit designations

- The data matches the expectations for any standard it claims to be meeting. If it's supposed to be a LOINC code it is in the format used for LOINC codes

> In keeping with the general approach of not cross checking data at its source, there is no proposal to check that it is, in fact, the correct LOINC code for the type of data presented. If the tester is knowledgeable enough to notice inconsistencies in the values provided they should be flagged, but manual cross-checking is not expected.

Logging and Classifying Errors

The tester should keep a log of all errors found. This log should include the row #, the pertinent field, and the type of error at a minimum. Additional information may be tracked if deemed useful.

If an error is found in a row randomly selected for spot checking, the tester should also examine a minimum of two rows above and two rows below it to look for the same error in the same field (other errors in those rows should be noted and addressed if noticed but they do not need to be sought out). If found in one of these rows, the error should be noted in a list of repeated problems. If found in two or more of these rows, the error should be noted as a recurring problem. If the same issue appears in the list of repeated problems more than N times it, too, should be considered a recurring problem.

If an error does not also occur in any of the rows immediately above or below it, it should be considered a one-off error by default. However, an ongoing tally of the combination of field and type of error should be kept to look for recurring but intermittent problems. If the same error occurs in the same type of field in more than 10% of sampled rows it should be identified as a recurring issue and treated the same way as errors repeated in nearby rows. Similarly, if the same general error type (such as a typo) occurs in any field in more than 5% of sampled rows it should be identified as a recurring general issue and addressed as such. Any errors that are close to these thresholds but not over them should be compared to the list of repeated errors from the non-sampled rows examined when issues are found and considered for inclusion in the list of recurring problems if found there as well.

Amount of Data to Check

If a file contains less than 2000 rows:

- a minimum of the larger of 40 or 3% of rows should be checked

- these rows should be randomly selected throughout the file

- rows checked because of issues found in nearby rows do not count toward these requirements

If a file contains between 2000-10000 rows:

- the file should be (virtually) split into four quadrants

- a minimum of the larger of 25 or 2% of the rows in each quadrant should be randomly checked

- rows checked because of issues found in nearby rows do not count toward these requirements

If a file contains more than 10000 rows:

- the file should be (virtually) split into at least six equal sections

- a minimum of the larger of 50 or 1% of the rows in each section should be randomly checked

- you can adjust these numbers proportionally if more than six sections are used as long as the results are at least the number of rows outlined above are checked in the whole file.
  - you could split the file into eight sections and use a formula of checking the larger of 40 or 1%
  - you could split the file into ten sections and use a formula of checking the larger of 30 or 1%

- rows checked because of issues found in nearby rows do not count toward these requirements

Rewards for Good Data

If a provider has three versions of the same file in a row without any recurring issues or two versions in a row without any errors (maybe beyond a very small threshold for typos?) they can halve the amount of rows needed for spot checks.

When a provider halves their spot check quotas, they restart their counts over. If they have another three versions in a row without recurring issues or two versions in a row without errors they may halve the number of rows being checked again. A provider with six consecutive versions with no errors will only have to check 1/8 of the number of rows outlined in the spec while a provider with some random one-off errors here or there but six consecutive versions with no recurring issues will only have to check 1/4 of the number of rows outlined in the spec.

However, any version with recurring issues resets the spot check requirements back to the initial levels agreed upon by the spec. A single version with more than M one-off errors resets back one level; if the spot check requirements had been halved twice, they go back to only being halved once.

In addition, any errors or issues found by the payer automatically reset the spot check requirements back to the initial levels agreed upon by the spec

Addressing Errors

Errors and issues found during the spot check process should be addressed and the file checked again before approved for submission. In addition, recurring issues should be investigated and the file creation process should be adjusted to eliminate (if possible) or reduce the occurrence of these issues.

Addressing One-Off Errors

Any one-off errors that do not meet the threshold for recurring issue should be fixed before the file is approved for transfer to the requesting payer.

Addressing Repeated Errors

Repeated errors that do not meet the designated threshold for recurring issues may be treated as one-off issues if a provider wishes. However, if resources permit, it would be better to investigate them as if they did meet the recurring threshold errors using the process described below.

Addressing Recurring Issues

If the tester identifies one or more recurring issues while examining a file, the transfer process should be halted while the errors are fixed.

The expectation for these errors is that they occur in sufficiently large number across the data rows not examined by the tester that the related data must be redone (or at least checked) for the entire file unless it can be proven the issue is isolated to a particular portion of the file.

The provider has two possible courses of action:

1. Manually check every row of the file for these specific problems and fix them, then later figure out what happened and how to fix it moving forward

   or

2. Determine the cause of the issue and address it, then reproduce the relevant data the correct way.

The second option may or may not require touching every row of data in the file. For example, the issue may be that a single person who prepared the second 500 and the last 500 rows of a 4000 row file misunderstood how to enter data in a particular field. The other 3000 rows are unaffected, so only the 1000 rows touched by this particular preparer need to be redone. The solution in this case may be to fix those 1000 rows and provide updated training to that preparer so he or she doesn't make the same mistake again. Whether it is more efficient to do the analysis to figure this out or to just redo every row correctly (or even if this can be determined without examining all 4000 rows in detail) is left to the determination of the provider.

Error Handling and Fault Tolerance

In an ideal world there would be strict error handling, but not everyone will adopt this right away. Instead, we will determine some rules and guidelines that impose some error handling and fault tolerance expectations but that allow for some flexibility in what payers will tolerate and accept.

To clarify, error handling is about identifying any specific problems with the files and data while fault tolerance is about how many and what type of errors to accept and absorb (as opposed to rejecting either specific rows or files).

This section is primarily concerned with issues in the files containing the requested patient data but will also address dealing with issues in the control files sent by payers here.

In version 1.0.0 of the MHDC HEDIS specification, validation rules are recommended guidelines. Participating organizations are not required to strictly enforce validation or mark rejected data in any way. The recommended deadlines for file processing or for resending corrected data were not hashed out before this decision was made and thus no specific timeframes are included in this specification even as guidelines. Any such deadlines are between each payer-provider pair.

File Rejection

The following errors should cause the entire file to be rejected immediately:

- bad file name
- bad file format
- wrong encoding
- wrong end of line characters
- data not in English
- bad delimiter
- deadline dates that do not match request deadlines
- inconsistent deadline dates in file names
- a required field (as defined by the relevant valid values file) is not present in any row in the file
- XXX

The following errors should cause the entire file to be rejected if they occur with the indicated frequency:

- any error in more than 1% of ID fields
- wrong data type for any field in more than 1% of that field
- wrong number of delimiters in more than 5% of rows
- data out of range for any field in more than 5% of that field
- any specific error in more than 10% of rows
- XXX

Row Rejections

The following errors should cause the entire row to be rejected:

- wrong number of delimiters in the row
- ID does not match a known patient
- missing required fields
- wrong data type for any field
- data out of range for any field

- unexpected end of line characters

- any field explicitly containing null

- malformed empty fields

- malformed enums

- malformed booleans

- invalid dates

- any field with a disallowed enum value

- any units field with a disallowed units designator

- units that do not match the data value provided

- data that clearly doesn't match what it's supposed to be (if noticed)

- XXX

Rejected rows should be marked as rejected by prepending a field containing R to the row.

Fixable Issues

The following issues may be fixed if noticed:

- infrequent simple typos where the intent was clear

- XXX

Ignorable Issues

The following issues should be ignored:

- extraneous content outside of the table format

- valid rows that contain data that wasn't requested (ex: cbc results when only a1c was requested)

- XXX

> If correctly formatted extra data is being sent regularly, mention it to the provider to save them the extra work in future data sets

Dealing with Rejected Data

All errors and row rejections should be logged. This log should include the row #, the pertinent field (if field specific), and the type of error at a minimum. Additional information may be tracked if deemed useful.

In addition, the type of error and the field it occurred in should be kept in a running tally for each file to track fault tolerance levels and determine whether the file needs to be rejected wholesale.

Error Handling in Control Files

This material will be added in a subsequent version of the MHDC HEDIS specification. It was not

hashed out prior to the decision not to enforce error handling rules in version 1.0.0 of the specification.

Data Dictionary

HEDIS information is sent by each provider to each payer monthly by secure FTP in the format described in this documentation.

Data Structure

The HEDIS data is currently split into groups of related content sent in separate flat text files containing rows of data each representing a specific instance of that type of data associated with a specific patient.

Data Rules

The HEDIS data specification includes rules and guidelines on how to represent different types of data.

Representing Patients with Patient IDs

Patient IDs uniquely identify the patient whose data is being sent in each row of the HEDIS data exchange. Patient IDs in version 1.0.0 of the MHDC HEDIS specification use a slightly modified form of a legacy ID format devised by Blue Cross.

This format may be revisited in the future to remove as much personally identifiable information as possible or for other reasons.

ID Format Rules

The current ID format has four components separated by hyphens. Each component may contain alphabetic characters, numbers, or hyphens as indicated below. No other characters are allowed.

Table 3. Patient ID components

Position	Name	Allowed Characters	Description
1	Member ID	Alphanumeric	The payer provided member ID for this patient. This consists of the patient's subscriber ID concatenated with the patient's two digit dependent code with no punctuation added between the two values.
3	Gender	Alphabetic	The single character code indicating the gender of the patient. Allowed options are M,F, and O for male, female, and other respectively.

Position	Name	Allowed Characters	Description
4	Date of Birth	Numeric with Hyphens	The patient's date of birth in the standard date format, YYYY-MM-DD
5	Partial First Name	Alphabetic	The first five characters of the patient's first name. All letters should be capitalized. If the patient's name is less than five characters long, the full name should be supplied and any remaining character slots omitted.

Examples

Sample lines from memberships data group:

```
045695946s953301-F-1970-03-27-KATHE | 1 | Smith | Katherine | ...
045645433f245700-M-1956-11-11-TIM | 1 | Corcoran| Tim | ...
```

Sample lines from observations data group:

```
045695946s953301-F-1970-03-27-KATHE | 2019-08-27 | 1234567890 | 56 |...
045695946s953301-F-1970-03-27-KATHE | 2019-08-27 | 1234567890 | 163 |...
045695946s953301-F-1970-03-27-KATHE | 2019-08-27 | 1234567890 | 89 |...
```

Other data groups will be similar.

Representing Numbers

The following rules and guidelines should be used when sending numerical data through the HEDIS data exchange:

- All numerical values should use base 10

- All numerical values should be included in Arabic numerals without scientific notation, punctuation (other than a decimal point if needed), or abbreviation.

  Use 1000 not 1,000 or 1.000 or 1K or 10^3 or 10^3 or M.

- Numbers are assumed to be floating point values unless otherwise restricted in specific field or unit descriptions

- Always use 0 for zero. -0 is not considered a valid value.

- Always include a leading 0 before the decimal point in numbers between 0 and 1

  Use 0.1 not .1.

- Only significant digits are included and precision is not tracked by trailing decimal zeros.

  Use 0.1 not 0.10.

- Limit values to no more than two significant decimal places unless there is a meaningful reason not to do so

  Values like 0.00056 are allowed but values like 0.100056 should be shortened to 0.1.

Representing Booleans

The following rules and guidelines should be used when sending boolean data through the HEDIS data exchange:

- Allowed values are true|false

  Do not use t|f or 1|0 or True|False

- Names of boolean properties should start with is, has, can, should or similar verbs

  Use isEnabled not enabled, hasCancer not cancer, canUpdate not updateable.

- Default is assumed to be unknown unless otherwise stated

  If not supplied, no inference about the value can be inferred

  If a particular boolean should be true or false by default the data dictionary will state so. The name of any such field should match the expected default value

  For example, in a region where most patients speak English, a boolean property indicating whether a particular patient does, in fact, speak English could be set to be either true or false by default if you want to assume the patient speaks English unless explicitly told otherwise. A boolean property that is true by default should use a name like isEnglishSpeaker while a similar property that is false by default should use something like cannotSpeakEnglish instead.

Representing Enumerated Values

The following rules and guidelines should be used when sending enumerated values through the HEDIS data exchange:

- Enumerated values should use complete words (no abbreviations and acronyms)
- Enumerated values should be a single term string (no spaces, dashes, underscores or other punctuation)
- Enumerated values should use camelCase for their names (nativeAmerican, pacificIslander, male, other)
- The allowed enumerated values will be listed in the description of each relevant field in alphabetical order (allowed unit values will be listed separately and referenced in the relevant field descriptions)

- If the value is unknown, leave the field blank unless directed otherwise in a specific field definition

- Exceptions can be made but must be explicitly noted and approved by the group

Representing Dates

The following rules and guidelines should be used when sending dates through the HEDIS data exchange:

- All dates should use the standard date format, YYYY-MM-DD

  Single digit months and single digit days should be represented as 0N. Use 2014-11-04 not 2014-11-4, 2016-09-12 not 2016-9-12, 2019-04-06 not 2019-4-6 or 2019-04-6 or 2019-4-06

- Some dates may not be known exactly.

  The portion of the date that is known should be supplied; the rest of the date should be omitted. For example:

  - If only a year is known, use YYYY to represent the date
  - If only a year and month is known, use YYYY-MM to represent the date

    The MHDC Data Governance Collaborative considered setting specific rules for exact dates to use when the exact date is unknown, but eventually decided against doing so because there are cases where time range requirements for {project} calculations might be violated if we assume the latest possible date but also cases where assuming the earliest possible date is disadvantageous.

Representing Units

The following rules and guidelines should be used when dealing with units in the HEDIS data exchange:

- The units field is always required for observations and laboratory tests
- Each type of measurement will have one and only one valid unit type

    If a test has more than one type of measurement each type of measurement will be treated as a different unit

  - All provider data must be converted to this unit prior to submission
  - All unit information must be stripped from the value and presented only in the matching units field.
- Each unit will have a single unit designator that represents that unit in the data. This is the value submitted in any units fields throughout the data set.
- Imperial/English/American units are favored over metric/SI units (mks or cgs) when both are common.

For example, internal height measurements of 5' 10" must be converted to a value field of 70 and a units field of inches.

Representing Trace Amounts

In some cases the result of a test that normally provides numerical results may come back as containing trace amounts or some other designation of a value too small to measure. These values should be represented as -1 and associated with the normal unit designator of the test. All payers are expected to understand the meaning of this value and convert it to their internal representation of trace values rather than rejecting the value as out of range.

> If any fields are added that allow negative values as valid values this will need to be adjusted. If such a case occurs, it should immediately be brought to the attention of MHDC.

Representing Unit Designators

The following rules and guidelines should be used when dealing with unit designators in the HEDIS data exchange:

- Unit designators typically use complete words unless they are formulas comprised of multiple atomic units (such as m/s)

- Unit designators use camelCase unless the term itself typically uses an initial capital letter (such as Fahrenheit)

- Unit designators with formulas of atomic units should use flat ASCII text representations of the formulas (use m/s^2 not m/s^2 and not a horizontal line divider between the nominator and denominator of the unit)

- Unit designators must match the specified value below exactly

> If additional units and unit designators are needed to adequately represent the data being exchanged, they should be negotiated through MHDC and added to the specification prior to use. Freeform values for units designators are not permitted.

Allowed Values

The following units and their corresponding unit designators are officially supported by MHDC:

Table 4. Supported Units

Unit Name	Unit Designator	Notes
general length	inches	Used for length measurements not specifically covered by other unit entries
general duration	seconds	Used for duration (time of a specific activity) measurements not specifically covered by other unit entries

Unit Name	Unit Designator	Notes
general weight	pounds	Used for weight measurements not specifically covered by other unit entries
general volume	gallons	Used for volume measurements not specifically covered by other unit entries
A1c result	percent	The average amount of glucose found in a patient's blood
albumin result	mg/dL	The amount of albumin found in a patient's urine
body height	inches	Used for height measurements taken during medical visits
body mass index (BMI)	kg/m^2	The calculated BMI value. This should be used for most adult BMI measurements. In keeping with using basic ASCII characters for unit designators, the caret character is used rather than an actual superscripted 2 (i.e. kg/m^2 is wrong).
body mass index percentile	percentile	BMI as a percentile measured against a historical reference population of the same age. This should be used primarily for BMI measurements of juvenile and adolescent patients. This should be paired with a value that's an integer between 0 and 100.
body temperature	Fahrenheit	Used for body temperature measurements taken during medical visits. Temperature readings taken in degrees Celsius should be converted to degrees Fahrenheit
body weight	pounds	Used for weight measurements taken during medical visits
diastolic blood pressure	mmHg	Used for the diastolic portion (bottom number) of a blood pressure reading taken during medical visits
hematocrit result	percent	The result of a hematocrit screening for anemia. This should be paired with a value that's a number between 0 and 100.

Unit Name	Unit Designator	Notes
hemoglobin result	g/dL	The result of a hemaglobin screening for anemia
LDL cholesterol level result	mg/dL	The results of an LDL cholesterol screening test or the LDL portion of a lipid screening panel.
lead test result	microg/dL	The results of a lead screening test in micrograms per deciliter. In keeping with using basic ASCII characters for unit designators, microg is used instead of \mug.
oxygen level	percent	The percentage of oxygen saturation in a patient's blood as read from a pulse oximeter during medical visits. This should be paired with a value that's a number between 0 and 100.
pulse	beatsPerMinute	The number of heartbeats per minute as recorded during medical visits
systolic blood pressure	mmHg	The systolic portion (top number) of a blood pressure reading taken during medical visits

Data Groups

The HEDIS data is organized in the following groups:

- diagnoses
- exceptions
- immunizations
- laboratoryServices
- medications
- memberships
- observations
- procedures

This data dictionary will describe the expected fields for each data group, identify its field type, list any constraints on those fields, and provide additional notes and comments on fields as needed. Each constraint is independent; just because a field has a minimum value does not mean the field is required to have a value present, just that if a value is supplied it must not be less than that minimum.

Each data group is an independent entity sent to payers as a separate file. The groups are listed alphabetically to make specific information easier to find.

While the data groups are independent, the fields within each group are not and must be included in the related data file in the order presented within the documentation.

Every data group contains a patient ID as its first field. This acts as a primary key that links patient-specific data together across the entire data domain. At the current time this identifier contains personally identifiable data about the patient it references; future versions of the specification may consider using a more standard format that limits the use of personal data as much as possible.

> ℹ️ Future versions of the specification may prepend a row number field to each data group as its first field to help log and fix data errors when stricter validation is enforced.

The diagnoses Data Group

The diagnoses data group defines patient diagnosis data.

Fields

The diagnoses data group includes the following fields:

Table 5. diagnoses Fields

Field Name	Field Type	Constraints	Description
patientId	string	required if any data is present	The identifier used to map data to a specific patient (i.e., a specific member of the insurance plan). This field is defined in every data group. See Representing Patients with Patient IDs for more information about IDs.
diagnosisCode	string	must be a valid code for the code system specified in the codeSystem field.	A code representing the diagnosis being reported in the current record.
codeSystem	enumerated string	The allowed values are: • icd9 • icd10 • snomed	The code system used for the supplied diagnosis in the diagnosisCode field.
status	enumerated string	The allowed values are: • active • inactive	Indicates whether the diagnosis is currently applicable or no longer applies.

Field Name	Field Type	Constraints	Description
startDate	date	date must be in the past; date must be before endDate if one exists.	The date the diagnosis started. The date must be in the standard date format, YYYY-MM-DD. This could map to a field containing the onset date in some provider systems.
endDate	date	date must be in the past; date must be after startDate if one exists.	The date the diagnosis ended. The date must be in the standard date format, YYYY-MM-DD. This could map to a field containing the abatement date in some provider systems.
reportedDate	date	date must be in the past	The date the diagnosis was first reported. The date must be in the standard date format, YYYY-MM-DD.
source	enumerated string	The allowed options are: • billing • encounter • problemList • thirdParty	Indicates where the diagnosis originated.

Examples

The following rows conform to the general structure of the diagnoses data group. However, they may or may not be valid for any given payer-provider pair depending on the valid values file in use.

```
045695946s953301-F-1970-03-27-KATHE|276.8|icd9|inactive|2014-08-12|2019-05-
06||encounter
045695946s953301-F-1970-03-27-KATHE|K52.9|icd10|active|2019-10-24|||thirdParty
045645433f245700-M-1956-11-11-TIM|E11.9|icd10|active|2018-10-31||2018-10-31|encounter
```

The exceptions Data Group

The exceptions data group defines patient-specific exceptions and exemptions.

Fields

The exceptions data group includes the following fields:

Table 6. exceptions Fields

Field Name	Field Type	Constraints	Description
patientId	string	required if any data is present	The identifier used to map data to a specific patient (i.e., a specific member of the insurance plan). This field is defined in every data group. See Representing Patients with Patient IDs for more information about IDs.

Field Name	Field Type	Constraints	Description
measureCode	enumerated string	The allowed options are: • PRC-BCS (Breast Cancer Screening) • PRC-CCS (Cervical Cancer Screening) • PRC-CDC-A1C (Diabetes Management: HbA1c Testing) • PRC-CDC-EYE (Diabetes Management: Eye Exam) • PRC-CDC-NPH (Diabetes Management: Medical Attention for Nephropathy) • PRC-CHL (Chlamydia Screening) • PRC-COL (Colorectal Cancer Screening) • OCM-CVD-LDL (Cardiovascular: Lipid Control) • OCM-CDC-A1C (Diabetes Management: HbA1c Control) • OCM-CDC-BP (Diabetes Management: Blood Pressure Control) • OCM-CDC-LDL (Diabetes Management: Lipid Control) • OCM-HYP-BP (Hypertension Management: Blood Pressure Control)	The measureCode indicates which screenings or care management related to a HEDIS measure is affected by the exception. The code consists of three segments, the first two required and the third optional. The first segment indicates whether the affected item is a process or an outcome (denoted by PRC or OCM respectively). The second segment identifies the relevant top-level measure using a three character code and the third segment identifies a specific submeasure using another three character code. Some measures may have associated processes, some may have associated outcomes, and some may have both. For example, diabetes management is a top-level measure with associated submeasures that has both specific processes and outcomes; A1c testing is a process denoted by PRC-CDC-A1C (process; clinical diabetes care; A1c) while A1c management is a submeasure denoted by OCM-CDC-A1C (outcome; clinical diabetes care; A1c).

Field Name	Field Type	Constraints	Description
exceptionCode	string with preferred values	The preferred values are: • age • medicalHistory • medicallyNotPossible • notAffordable • notApplicable • postponed	For the MHDC HEDIS version 1.0.0 specification the preferred values should be used if applicable but other values may also be used; this is not a strict enumeration. However, if you find another value is necessary, please report it to MHDC so it can be added to the preferred list and included when this preferred list becomes an enforced enumerated list in the future.
reason	string	This field is expected to be significantly longer than the typical string value; this should be accounted for when consuming the data so no information is lost.	A freeform description of the exception and why the exception is valid expanding on the provided exception code. For example, if a patient needs to get a colonoscopy every two years because of a combination of frequent abnormal benign polyps and unexplained gastrointestinal symptoms, a valid reason field might be: Patient has exhibited numerous hamartomatous polyps on each of his last three colonoscopies. Further, patient has complained of frequent nausea, constipation, and diarrhea without known cause. Patient has declined genetic testing; colonoscopies should be performed every two years as long as hamartomatous polyps and symptoms persist. If polyps persist without symptoms, colonoscopy schedule may be readjusted to every three years. If no polyps of any kind are found in two consecutive colonoscopies, colonoscopy schedule may be readjusted to every five years.

Field Name	Field Type	Constraints	Description
frequency	string	none	Indicates a change in frequency rather than an exception to including data at all. For example, some patients have medical conditions that warrant having colonoscopies every 2, 3, or 5 years rather than every 10 years. These exceptions to the normal expectations would be noted here by supplying a frequency value for colonoscopies.
startDate	date	date must be in the past; date must be before endDate if one exists.	The date the exception started. The date must be in the standard date format, YYYY-MM-DD.
endDate	date	date must be in the past; date must be after startDate if one exists.	The date the exception ended. The date must be in the standard date format, YYYY-MM-DD.

Examples

The following rows conform to the general structure of the exceptions data group. However, they may or may not be valid for any given payer-provider pair depending on the valid values file in use.

```
045645433f245700-M-1956-11-11-TIM|PRC-COL|medicalHistory|Patient has exhibited
numerous hamartomatous polyps on each of his last three colonoscopies. Further,
patient has complained of frequent nausea, constipation, and diarrhea without known
cause. Patient has declined genetic testing; colonoscopies should be performed every
two years as long as hamartomatous polyps and symptoms persist. If polyps persist
without symptoms, colonoscopy schedule may be readjusted to every three years. If no
polyps of any kind are found in two consecutive colonoscopies, colonoscopy schedule
may be readjusted to every five years.|once every two years|2009-10-10|
045695946s953301-F-1970-03-27-KATHE|PRC-CHL|notApplicable|Patient not sexually active;
one year waiver renewable if still pertinent|2018-10-01|2019-09-30
045695946s953301-F-1970-03-27-KATHE|PRC-CHL|notApplicable|Patient not sexually active;
one year waiver renewable if still pertinent|2019-10-01|2020-09-30
```

The immunizations Data Group

The immunizations data group defines patient immunization data.

Fields

The immunizations data group includes the following fields:

Table 7. immunizations Fields

Field Name	Field Type	Constraints	Description
patientId	string	required if any data is present	The identifier used to map data to a specific patient (i.e., a specific member of the insurance plan). This field is defined in every data group. See Representing Patients with Patient IDs for more information about IDs.
serviceProviderNpi	integer	minimum value= 1000000000. Maximum value= 29999999999	The NPI number of the service provider.
serviceDate	date	date must be in the past	The date the service was performed. The date must be in the standard date format, YYYY-MM-DD. This could map to a field containing the administered date in some provider systems.
immunizationCode	string	must be a valid code for the code system specified in the codeSystem field.	A code representing the immunization being reported in the current record.
codeSystem	enumerated string	The allowed values are: • cpt • cvx • mapped • snomed	The code system used for the supplied immunization in the immunizationCode field. If mapped is used, it means that a freeform value will be supplied in the corresponding immunizationCode field; this value will not conform to any of the standard code systems supported by the MHDC HEDIS specification.
immunizationName	string	none	A freeform string representing the common name of the immunization.

Examples

The following rows conform to the general structure of the immunizations data group. However, they may or may not be valid for any given payer-provider pair depending on the valid values file in use.

```
045695946s953301-F-1970-03-27-KATHE|24466897751|2018-09-18|88|cvx|influenza vaccine
045645433f245700-M-1956-11-11-TIM|25673453483|2018-11-03|86198006|snomed|flu vaccine
045695946s953301-F-1970-03-27-KATHE|24466897751|2018-10-04|88|cvx|influenza vaccine
045645433f245700-M-1956-11-11-TIM|25673453483|2018-11-03|12866006|snomed|pneumonia
vaccine
045645433f245700-M-1956-11-11-TIM|25673453483|2019-09-15|tetanus|mapped|
```

The laboratoryServices Data Group

The laboratoryServices data group defines patient laboratory test results data.

Fields

The laboratoryServices data group includes the following fields:

Table 8. laboratoryServices Fields

Field Name	Field Type	Constraints	Description
patientId	string	required if any data is present	The identifier used to map data to a specific patient (i.e., a specific member of the insurance plan). This field is defined in every data group. See Representing Patients with Patient IDs for more information about IDs.
orderingProviderNpi	integer	minimum value= 1000000000. Maximum value= 29999999999	The NPI number of the ordering provider.
labId	string	none	An identifier for the laboratory analyzing the test and providing the results. In version 1.0.0 use the freeform name of the laboratory. MHDC plans to standardize this field in future versions, possibly using the CLIA code preferred by the CDC.
serviceCode	string	must be a valid code for the code system specified in the codeSystem field.	A code representing the service being reported in the current record.

Field Name	Field Type	Constraints	Description
codeSystem	enumerated string	The allowed values are: • cpt • hcpcs • loinc • mapped • snomed	The code system used for the supplied service in the serviceCode field. If mapped is used, it means that a freeform value will be supplied in the corresponding serviceCode field; this value will not conform to any of the standard code systems supported by the MHDC HEDIS specification.
serviceDate	date	date must be in the past	The date the service was performed. The date must be in the standard date format, YYYY-MM-DD. This could map to a field containing the collection date in some provider systems.
resultDate	date	date must be in the past; date must be after serviceDate if one exists	The date the results were processed by the laboratory. The date must be in the standard date format, YYYY-MM-DD.
result	number	must be -1 or greater or equal to 0	The numerical result of a specific test in the units system specified in the units field. A value of -1 indicates that trace amounts of the substance being measured are present. Results that cannot be represented as a single number in an allowed unit should be reported using the summaryOfFindings field which supports more verbose and freeform results. No test results captured in version 1.0.0 of the MHDC HEDIS specification include negative numbers as valid results; if this is expected to change in the future, contact MHDC immediately to start discussion of handling this case in future versions.

Field Name	Field Type	Constraints	Description
unit	enumerated string	See Allowed Values in the section on representing units for information about the allowed unit designators and what they mean.	The unit designator for the service. Only one unit is valid per type of measurement; you may need to convert your source data to report it in the specified unit. If the desired type of unit is not represented in the list of allowed values, contact MHDC to have it added to the specification.
summaryOfFindings	string	This field is expected to be significantly longer than the typical string value; this should be accounted for when consuming the data so no information is lost.	Provides a summary of a collection of findings rather than the results of an individual test or measurement when such a summary is appropriate. For example, while individual blood chemistry results should be sent as separate laboratory entries as appropriate with each test result as a single piece of data stored in the result/unit field pair, discussions of the complete chemistry panel and its implications might be sent separately as a summary of findings.

Examples

The following rows conform to the general structure of the laboratoryServices data group. However, they may or may not be valid for any given payer-provider pair depending on the valid values file in use.

```
045645433f245700-M-1956-11-11-TIM|12435687090|22D0075012|4548-4|loinc|2018-12-12|2018-
12-14|7.8|percent|
045695946s953301-F-1970-03-27-KATHE|24466897751|22D0073159|102739008|snomed|2018-12-
20|2018-12-29|88|mg/dL|
045645433f245700-M-1956-11-11-TIM|12435687090|22D0075012|4548-4|loinc|2019-03-28|2019-
03-29|7.3|percent|
045645433f245700-M-1956-11-11-TIM|12435687090|22D0075012|4548-4|loinc|2019-06-30|2019-
07-03|6.8|percent|
045645433f245700-M-1956-11-11-TIM|12435687090|22D0075012|4548-4|loinc|2019-10-04|2019-
10-05|7.1|percent|
045695946s953301-F-1970-03-27-KATHE|26363522106|22D0074386|G0105|hcpcs|2018-12-
20|2018-12-29|||two polyps removed during procedure analyzed at lab; one 2.22mm in
diameter and 2.54mm in diameter. Both benign.
045645433f245700-M-1956-11-11-TIM|12435687090|22D0075012|4548-4|loinc|2019-12-22|2019-
12-27|7.0|percent|
```

The medications Data Group

The medications data group defines data related to the medications a patient has taken during the year including current medications.

Fields

The medications data group includes the following fields:

Table 9. medications Fields

Field Name	Field Type	Constraints	Description
patientId	string	required if any data is present	The identifier used to map data to a specific patient (i.e., a specific member of the insurance plan). This field is defined in every data group. See Representing Patients with Patient IDs for more information about IDs.
orderingProviderNpi	integer	minimum value= 1000000000. Maximum value= 29999999999	The NPI number of the ordering provider.
orderDate	date	date must be in the past; date must be before startDate if one exists	The date the medication was ordered. If this record is for a refill, this should be the date the refill was ordered and not the date of the initial order for the medication. The date must be in the standard date format, YYYY-MM-DD.
startDate	date	date must be in the past; date must be before endDate if one exists.	The date the medication started. The date must be in the standard date format, YYYY-MM-DD.
endDate	date	date must be in the past; date must be after startDate if one exists.	The date the medication ended. The date must be in the standard date format, YYYY-MM-DD.

Field Name	Field Type	Constraints	Description
status	enumerated string	The allowed values are: • active • continued • discontinued • increased • patientStopped • refilled • telephoneStart	Indicates whether the medication is currently applicable or no longer applies.
medicationCode	string	must be a valid code for the code system specified in the codeSystem field.	A code representing the medication being reported in the current record.
codeSystem	enumerated string	The allowed values are: • rxNorm • ndc	The code system used for the supplied medication in the medicationCode field.
medicationName	string	none	A freeform string representing the common name of the medication.
supplyLength	integer	minimum value=1	The number of days of medication included in each prescription fill
quantity	integer	minimum value=1	The total number of pills, shots, or other medication included in each prescription fill
frequency	string	none	How often the medication is taken by the patient.
dose	number	must be a positive number	The number of pills, puffs, injections, spoonfuls, etc. in each dose of the medication (i.e. how many items are taken at one time). Supports values like 0.5 to indicate taking half of a pill per dose.

Field Name	Field Type	Constraints	Description
strength	string	none	The amount of medication in each item taken by the patient in milligrams or other appropriate unit. The strength should include both the value and the corresponding unit of measurement. The strength is per item, not per dose. If the patient takes 2 10mg pills per dose, the strength should be listed as 10mg. If the patient takes half of a 1mg pill per dose, the strength should be listed as 1mg. ℹ Future versions of the MHDC HEDIS specification may split strength into two fields; one with the value and one with the unit to be more consistent with how we handle values with units generally.
deliveryMechanism	string	none	The way the medication is taken: pill, liquid, ointment, nasal inhaler, oral inhaler, injection, auto-injector, etc. ℹ Future versions of the the MHDC HEDIS specification should consider turning deliveryMechanism into an enumerated field

Examples

The following rows conform to the general structure of the medications data group. However, they may or may not be valid for any given payer-provider pair depending on the valid values file in use.

```
045695946s953301-F-1970-03-27-KATHE|12323434567|2017-03-
13|||active|861007|rxNorm|metformin|30|60|twice daily|1|500 mg|pill
045695946s953301-F-1970-03-27-KATHE|12323434567|2018-03-13||2019-10-
02|discontinued|0093-7254|ndc|glimiperide|15|30|once daily|0.5|1 mg|pill
045695946s953301-F-1970-03-27-KATHE|12323434567|2019-10-02|||active|0093-
7254|ndc|trulicity|28|4|once weekly|1|1.5 mg/.5mL|auto-injector
```

The memberships Data Group

The memberships data group defines patient identification and demographic data.

Fields

The memberships data group includes the following fields in the order specified below:

Table 10. memberships Fields

Field Name	Field Type	Constraints	Description
patientId	string	required if any data is present	The identifier used to map data to a specific patient (i.e., a specific member of the insurance plan). This field is defined in every data group. See Representing Patients with Patient IDs for more information about IDs.
subscriberNumber	string	required	The subscriber ID assigned to this patient by the payer.
dependentCode	string	required. Must be two numeric characters long; leading 0 characters are acceptable.	Indicates the position of the patient in the family structure. Each payer may use its own scheme for assigning values. In general, the subscriber is assigned either 00 or 01 but this is not required. The combination of subscriberNumber and dependentCode values should uniquely identify the patient to the payer.
lastName	string	none	The last name of the patient. The name should use the casing, spacing, and punctuation of the actual name. Suffixes such as Jr., Sr., III should follow the alphabetic portion of the name with a single space between them.
firstName	string	none	The first name of the patient. The name should use the casing, spacing, and punctuation of the actual name. Prefixes such as titles should not be included.
middleInitial	string	maximum length=1	The middle initial of the patient. This letter should be capitalized and not followed by a period or any other punctuation.

Field Name	Field Type	Constraints	Description
gender	enumerated string	The allowed options are: • F • M • O	The current gender of the patient. The names of these values violate naming conventions per decision of the Data Governance Collaborative. O (for other) indicates a known gender that is not male or female. If the gender is unknown data for this field should be omitted. At the current time, there is no specific definition of what gender is and no guarantee that is is consistently defined across all data records. For any given record, gender could be gender identity, biological gender, birth gender, or any number of other variants. Future versions of the MHDC HEDIS specification may impose specific definitions on gender and may add additional gender fields and gender choices.
birthDate	date	date must be in the past	The date of birth of the patient in YYYY-MM-DD format.

Field Name	Field Type	Constraints	Description
phoneNumber	string	minimum ten characters	The phone number of the patient. This number may be an international phone number. If no country code is supplied, the number is assumed to be a standard ten digit phone number from the United States or Canada. In general punctuation and spaces should be stripped out of the number; the exception to this is for supplying country codes which should be presented using the following syntax: `+ code number` where *code* is a valid country code and *number* is the local number within that country.
spokenLanguage	string	none	The spoken language of the patient. If no language is supplied, the default language is English. If only one language is known for the patient, assume it is the spoken language of the patient.
writtenLanguage	string	none	The written language of the patient. If no written language is supplied, the default is to use the spokenLanguage value. If no language is supplied in either field, the default language is English.
race	enumerated string	The allowed options are: • africanAmerican • asian • hispanic • nativeAmerican • other • pacificIslander • white • unknown	The identified race of the patient. Note that an explicit value of unknown is supplied; if the race is unknown this value should be used rather than supplying a blank field.

Field Name	Field Type	Constraints	Description
hispanicEthnicity	enumerated string	The allowed options are: • isHispanic • isNotHispanic • refusedToSpecify • unknown	Indicates whether the patient self-identifies as Hispanic. This value should agree with the value of the isHispanic field that follows, but is slightly more expansive. Note that an explicit value of unknown is supplied; if the race is unknown this value should be used rather than supplying a blank field.
isHispanic	boolean	none	Indicates whether the patient is Hispanic. If not supplied, the value is assumed to be unknown. This value should not contradict the value of the earlier hispanicEthnicity field, but provides fewer options.

Examples

The following rows conform to the general structure of the memberships data group. However, they may or may not be valid for any given payer-provider pair depending on the valid values file in use.

```
045695946s953301-F-1970-03-27-KATHE|045695946s9533|01|Montgomery|Katherine|J|F|1970-
03-27|6175551212|||africanAmerican|isHispanic|true
045645433f245700-M-1956-11-11-TIM|045645433f2457|00|Fredricks Jr|Tim|O|M|1956-11-
11||||white|isNotHispanic|false
A67d88999BB56204-M-2001-03-24-VERNO|A67d88999BB562|04|Malick|Vernon|R|M|2001-03-
24|7815550045|||asian|isNotHispanic|false
A6755559HKK2502-O-1999-06-21-VERA|A6755559HKK25|02|Malnikova|Vera|N|O|1999-06-
21|9785553567|Russian||white|isNotHispanic|
```

The observations Data Group

The observations data group defines patient vital signs and other observational data taken during medical visits.

Fields

The observations data group includes the following fields:

Table 11. observations Fields

Field Name	Field Type	Constraints	Description
patientId	string	required if any data is present	The identifier used to map data to a specific patient (i.e., a specific member of the insurance plan). This field is defined in every data group. See Representing Patients with Patient IDs for more information about IDs.
serviceProviderNpi	integer	minimum value= 1000000000. Maximum value= 29999999999	The NPI number of the service provider.
serviceDate	date	date must be in the past	The date the service was performed. The date must be in the standard date format, YYYY-MM-DD.
observationCode	string	must be a valid code for the code system specified in the codeSystem field.	A code representing the observation being reported in the current record.
codeSystem	enumerated string	The allowed values are: • cpt • cpt2 • hcpcs • loinc • mapped • snomed	The code system used for the supplied observation in the observationCode field. If mapped is used, it means that a freeform value will be supplied in the corresponding observationCode field; this value will not conform to any of the standard code systems supported by the MHDC HEDIS specification. The code system selected for each record must adequately describe the unique observation supplied for that record. For example, the code system must explicitly distinguish between diastolic and systolic blood pressure and allow providers to explicitly indicate the record provides one or the other. Systems that only supply a code for blood pressure generally may not be used to supply blood pressure results.
value	number	none	The numerical value of the vital sign or observation in the units indicated by the unit field.

Field Name	Field Type	Constraints	Description
unit	enumerated string	See Allowed Values in the section on representing units for information about the allowed unit designators and what they mean.	The unit designator for the observation. Only one unit is valid per type of measurement; you may need to convert your source data to report it in the specified unit. If the desired type of unit is not represented in the list of allowed values, contact MHDC to have it added to the specification.

Examples

The following rows conform to the general structure of the observations data group. However, they may or may not be valid for any given payer-provider pair depending on the valid values file in use.

```
045645433f245700-M-1956-11-11-TIM|12435687090|2019-05-23|8480-6|loinc|142|mmHg
045645433f245700-M-1956-11-11-TIM|12435687090|2019-05-23|8462-4|loinc|82|mmHg
045645433f245700-M-1956-11-11-TIM|12435687090|2019-05-23|162763007|snomed|188|pounds
045645433f245700-M-1956-11-11-TIM|12435687090|2019-05-
23|386725007|snomed|97.7|Fahrenheit
```

The procedures Data Group

The procedures data group defines data related to procedures the patient has undergone throughout the year.

Fields

The procedures data group includes the following fields:

Table 12. procedures Fields

Field Name	Field Type	Constraints	Description
patientId	string	required if any data is present	The identifier used to map data to a specific patient (i.e., a specific member of the insurance plan). This field is defined in every data group. See Representing Patients with Patient IDs for more information about IDs.
orderingProviderNpi	integer	minimum value= 1000000000. Maximum value= 29999999999	The NPI number of the ordering provider.

Field Name	Field Type	Constraints	Description
serviceProviderNpi	integer	minimum value= 1000000000. Maximum value= 29999999999	The NPI number of the service provider.
procedureCode	string	must be a valid code for the code system specified in the codeSystem field.	A code representing the procedure being reported in the current record.
codeSystem	enumerated string	The allowed values are: • cpt • cpt2 • hcpcs • icd9 • icd10 • mapped • snomed	The code system used for the supplied procedure in the procedureCode field. If mapped is used, it means that a freeform value will be supplied in the corresponding procedureCode field; this value will not conform to any of the standard code systems supported by the MHDC HEDIS specification.
serviceDate	date	date must be in the past	The date the service was performed. The date must be in the standard date format, YYYY-MM-DD.
results	string	This field is expected to be significantly longer than the typical string value; this should be accounted for when consuming the data so no information is lost.	A freeform string indicating the contents of the procedure report or other pertinent results from the procedure itself. This is not meant to include results from biopsies or other laboratory services performed later even if they are performed on tissue or other substances removed during the procedure; those should be recorded as laboratoryServices records.
source	enumerated string	The allowed options are: • billing • encounter • problemList • thirdParty	Indicates where the procedure originated.

Examples

The following rows conform to the general structure of the procedures data group. However, they may or may not be valid for any given payer-provider pair depending on the valid values file in use.

```
045645433f245700-M-1956-11-11-TIM|12120045645|12004744127|G0105|hcpcs|2019-11-
11||encounter
045695946s953301-F-1970-03-27-KATHE|20001203055|11114680257|90226004|snomed|2015-05-
14|Pap smear normal|encounter
045695946s953301-F-1970-03-27-KATHE|20001203055|24956700021|58570|cpt|2018-06-
06|Uterus removed as precaution based on genetic screening. Laproscopic procedure
successful.|thirdParty
```