Status of the MHDO Health Data Warehouse Master Indexes

*October 14, 2015*

# Overview

MHDO’s vision of the Health Data Warehouse project originally called for the creation of three master indexes: payer, patient, and provider. In its call for proposals, they stated that these indexes would “create a shared utility that will provide value for multiple entities through the state.” These indexes would be used across claim, hospital, and other data streams to provide “consistent, accurate and historical demographic data” on patients, providers, and payers.

During the past two years, MHDO and its contractor HSRI have made major strides towards meeting these requirements. This document outlines the additional steps that need to be taken to achieve the goals of the RFP and a proposed timeline for this work.

# Master Payer Index

The Mater Payer Index is intended to provide “consistent, accurate, historical, and current demographic data on the payers…reported across the claims, inpatient/outpatient, and other data streams. Each payer was to appear only once across all streams in this index.”

Currently, every payer that submits claims data is assigned an MHDO Assigned Code. This unique identifier and the payer name are stored on the Payer table, along with the date of addition. The MHDO Assigned Code and Payer name are distributed to data users as a part of the Data Release process. In addition, HSRI has created a report that documents the activation and deactivation dates of any new payers or payers that no longer meet the submission threshold.

The hospital does not make use of the MHDO Assigned Code to identify third party payers. Discharge data are submitted with payer identification numbers. According to Chapter 241, payers are identified with hospital electronic billing payer codes, NAIC codes, or MHDO individual payer codes. In addition to the payer number, the payer name is also included. During the ingestion process, each payer code is categorized into one of 10 categories, such as “self pay,” Medicaid,” etc.

## Challenges

One of the key challenges in creating a Payer Index that covers both the claims and the hospital data is that the information is submitted at different levels of granularity between the two sources. The claims data identifies all claims related to a given company with a single 5-digit code and a suffix character that has no standardized semantic meaning. The hospital data, on the other hand, has payer names that may denote specific plan types within a company, such as ‘Aetna HMO’, ‘Aetna PPO,’ Aetna Medicare,’ etc. Also the hospital includes payers that never appear in the claims data, such as workers comp, self-pay, and self-insured plans.

Another challenge is that the payer identification number cannot be relied upon to be an NAIC code. Thus, it is of limited use in relating hospital payer information to claim information.

## Proposed Solution

The payer index will contain to report payer information at the MHDO-Assigned code level. MHDO and HSRI are working with APCD data users to develop a shell for an enhanced table that will contain information about each company (status, parent company, NAIC, FEIN etc.), current business (data start/end dates, file types, business relationships etc.), and historical business (data start/end dates, file types, business relationships etc.).

During the hospital data ingestion process, in addition to assigning each unique payer value to one of 10 categories, each will be related to either a) an existing MHDO-Assigned Code or b) a special code indicating “Other Payer”. This will allow data users to more easily relate claims data aggregate results with discharge data aggregate results and provide a richer source of information than the current 10 categories.

## Proposed Timeline

|  |  |  |
| --- | --- | --- |
| Task | Start Date | End Date |
| Assign MHDO-Assigned Code to Historical Data | November 2015 | December 2015 |
| Begin Distributing Payer Index with APCD Requests | January 2016 | -- |
| Begin Distributing Payer Index with Hospital Data Requests | January 2016 | -- |
| Create Process to Assign MHDO-Assigned Code during Ingestion | April 2016 | July 2016 |

# Master Patient Index

The Master Patient Index is intended to provide “consistent, historical, and current demographic data on the patients reported across the claims, inpatient/outpatient, and other data streams.” Each patient in the index would receive a single unique identifier across all the streams.

Currently, the MHDO calculates a unique member ID based upon the member SSN, the subscriber SSN, and/or the contract number on the claim. A project was recently completed that performed partial de-duplication of historical member IDs. However, it is known that some duplication of IDs (that is, situation where one individual has more than one member ID) still exist due to ambiguities in the data.

In the hospital data, there is no member ID. Individuals may be able to be tracked across time within a facility by medical record number, however, mergers and system changes may prevent tracking even within a facility. In addition, the SSN requested in the Chapter 241 layout is in the third party payer loop and it may or may not be a patient SSN.

## Challenges

In both the claims and the hospital data, there are situations where we lack the identifiers necessary to create a good quality member ID. That is, if we only have DOB, gender and contract number in the claims data, we can only match this member across payers when the same contract number is elsewhere associated with an SSN. In the hospital data, we would only relate this individual across facilities and to the claims data if the SSN on the payer loop happens to be the individual’s SSN (which it often appears is not the case).

The revised Chapter 243 already calls on individual payers to ensure that the same identifier is used between all claims and eligibility files for a given individual. However, this still leaves the problem of tracking an individual between payers.

## Proposed Solution

The MHDO will ask payers and submitters of hospital discharge data to provide a periodic demographics file. This file would include the identifiers they are currently using for an individual and additional demographic information that may allow matching between payers and facilities, such as SSN (the ideal case), patient name, address, and phone number.

This information will then be used to improve the assignment of the Unique Person Identifier on the claims side and allow its addition on the hospital side. The UPID would then be distributed alongside or in place of the current member identifiers in future data releases. Information from the patient index appropriate to the release level, such as county, age (pegged to a reference date), and gender would be distributed in a patient index release file to accompany data releases.

## Proposed Timeline

|  |  |  |
| --- | --- | --- |
| Task | Start Date | End Date |
| MHDO Requests Demographics Files | November 2015 | -- |
| Voluntary Demographics Files Submitted through Portal (provide a naming convention to bypass validation) | December 2015 | May 2016 |
| Demographics Files Ingested to Patient Index | January 2016 | June 2016 |
| Process for Assigning UPID with Demographics Files | March 2016 | June 2016 |
| Assign UPIDs with new Process to Historical Data | July 2016 | July 2016 |
| Begin Distributing UPID and Patient Index Release File | July 2016 | -- |

# Master Provider Index

The Master Provider Index is intended to provide “consistent, accurate, historical, and current demographic data on the medical providers reported across the claims, inpatient/outpatient, and other streams.” This index was required to include the National Provider Identifier (NPI) of the provider.

In the claims data, the NPI is the preferred source of provider identification. When an NPI is present on a record, it is assumed to correctly identify the individual. In addition, every unique combination of provider fields is assigned a PRVIDN. When possible, these PRVIDNs are matched to master file records by DPCID. These other provider fields include provider name, payer-assigned provider number, etc. In the data release files, the PRVIDNs and DPCIDs are provided to all recipients, allowing the recipients to identify claims related to a given provider. For recipients who receive provider identifiable information, the NPI and name are also provided.

In the hospital data, the provider is identified by an NPI and a name. No other provider information is available. This information is used to assign an MHDO-assigned “encrypted ID” that is then included in the data release. We do not currently distribute NPI to data requestors for the discharge data. The MHDO-assigned provider identifiers in the hospital data have no relation to the ones assigned in the claims data.

## Challenges

Since the claims and hospital data use different provider identifiers, these data cannot be easily linked by provider, especially since current data requestors do not receive any provider identifiable information for the hospital data.

## Proposed Solution

Since both the claims and the hospital data have high rates of population for the NPI, we propose that this also be the basis for matching providers between the two data streams. To this end, we propose that a DPCID be assigned to every hospital record. This DPCID will match to a Master Provider Index row for the same provider. Specialty and taxonomy codes will be those that appear in the NPPES registry.

The claims data process of assigning DPCID would be modified to point to this Master Provider Index rather than the current provider master file. Entries from the existing master file that lack an NPI entry would be loaded to this index as well, making it functionally equivalent to the current master files. The difference would be that the matching order would be changed to use the NPI before doing any “fingerprint matching” from the detail file.

## Proposed Timeline

|  |  |  |
| --- | --- | --- |
| Task | Start Date | End Date |
| Create new DPCIDs for Provider Index | November 2015 | December 2015 |
| Modify claims data logic to match NPI then “fingerprint” | December 2015 | -- |
| Add DPCID to Hospital Data | December 2015 |  |
| Add DPCID logic to Hospital Ingestion Process | July 2016 | -- |
| Begin distributing Master Provider Index release file with new DPCID | January 2016 |  |