Status of the MHDO Health Data Warehouse Master Indexes

*May 11, 2016*

# 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 three 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 fully achieve the goals of the RFP. This document is meant to guide the conversation between HSRI and MHDO to rigorously define the end state of this development (“define done”) and a produce an achievable 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.”

## APCD

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 current Payer Index that is distributed with the data releases and available on the website contains the MHDO-Assigned payer codes and payer names of all of the entities contained within the claims data. This is augmented by the Activation/Deactivation report that accompanies each release which details the MHDO-Assigned Payer Code, Payer Name, Activation, Deactivation, Start and End Dates, Expected Release for first inclusion, Data types, Medical, Medicare Medical Part C, Pharmacy, Medicare Pharmacy Part C and D, and Dental covered lives and submission frequency for all newly activated and deactivated payers.

## Hospital Encounter Data

The hospital does not make use of the MHDO Assigned Code to identify third party payers. Discharge data are submitted with payer names and 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 9 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, historically, the payer identification numbers have not been drawn from a standardized list. Since the number may be an electronic billing payer code, an NAIC code, or an MHDO individual payer code, this information can be very challenging to use and is of limited use in relating hospital payer information to claim information.

## Decisions to Be Made

There are several options on how we should proceed:

1. Is there any additional information that should be added to the Payer Index that is distributed with the APCD data? Are there any additional Payer reports that should be developed for distribution with data releases?
2. How precisely will the standardized list of hospital payers be developed that will be distributed to data submitters going forward? What are the roles and responsibilities for the maintenance and distribution of this list and how are home-grown codes handled?
3. Should a list of the hospital payers in historical data be integrated with the APCD Payer Index or should there be a separate Hospital Payer Index?

# 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.

## APCD

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.

## Hospital Encounter 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.

## Decisions to be Made

1. Should we request an additional Demographics file from APCD and/or Hospital data requestors to help in the creation of Member IDs?
2. Should a separate MemberID index be distributed with the claims and/or hospital data summarizing member information (E.g., MemberID, DOB, Gender, current County)?
3. Should the SSN on the payer loop of the hospital data be augmented with an SSN on the patient data loop? Currently, we cannot detect when the payer loop SSN is the patient or when it is another entity, which limits our ability to link hospital and claims data.

# 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.

## APCD

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 Provider Identifier (PRVIDN). When possible, these PRVIDNs are matched to master file records by an internally assigned Master Provider Identifier (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.

## Hospital Encounter Data

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 hospital encounter 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 operating and attending providers within the hospital data. Currently, they only receive identification of the facility itself, which is often (but not always) analogous to the billing provider in the claims data.

## Decisions to Be Made

1. Should the provider identifiers for attending and operating providers in the hospital data be the same as we provide in the claims data? This would allow for the distribution of NPI for data requestors who requested provider identified data and would potentially allow for the linking of attending and servicing providers between hospital and claims data.