MA APCD Release 8.0 Master Patient Index (MPI) Data Exclusion

Background

In MA APCD Release 8.0 CHIA used our Master Patient Index (MPI) approach to create unique MEMBERLINKEID for all members within the MA APCD. The MEMBERLINKEID provides a way to identify a member using a unique identifier that is not personally identifiable. The MEMBERLINKEID's provide the researchers' the ability to track members across time, sites of service, products, and payers.

A write-up of the new MA APCD Master Patient Index process can be found at:
http://www.chiamass.gov/assets/docs/p/apcd/Overview-of-New-MA-APCD-Master-Patient-Index.pdf

Input File Analysis for Data Exclusion

CHIA undertook a substantial analysis step, carefully reviewing all eligibility files for anomalies or duplications that would meaningfully decrease the utility of the MEMBERLINKEID. As MA APCD consists of submissions from both the risk holder (payer) and, in some cases, an administrator it was determined that in some circumstances this created a significant amount of duplication. In these cases the administrator data was excluded from the MPI process; in other cases some submitters' data was found to be of poor quality.

Duplication/poor quality data fell into three categories of MPI exclusion:
1. A submitter's entire eligibility data was excluded.
2. An individual year for a submitter, or a subset of years, were excluded.
3. A subset of individual records were excluded.

A full listing of these exclusions by OrgID is found in Attachment A of this document.The records for these excluded records will have no MEMBERLINKEID attached.

Conclusion

After CHIA tested the newly created MEMBERLINKEIDs, CHIA found significant improvement in the match rates and significant reduction in the number of "orphan" records. "Orphan" records are records with no connection to any claims or any other membership records.

Note: Since values for MEMBERLINKEID are recalculated prior to each new release, this surrogate key **cannot** be used to link data across any other prior or subsequent releases. CHIA has created a surrogate key mapping across releases. Researchers may request this crosswalk.

Attachment A: MPI Exclusion List

1.   Complete Exclusion of Submitters

The first type of data exclusion from the MPI process is the identification of specific data submitters that were chosen to not be included in the MPI process. These submitters were reviewed by CHIA and deemed that they would cause damage to the creation of the MEMBERLINKID. Two key issues were found. First, the data from an administrator was deemed to contain a high level of duplication to the risk holder. Second, was the discovery that throughout the data submissions many unique individuals were found to appear under a single CHIA Organization ID (OrgID)/Carrier Specific Unique Member ID (CSUMID). These submitters were excluded from the MPI process.

ORG ID
10942
10950
10953
10954
11069
11239
11264
11364
11371
11500

2.   Year-by-Year Exclusion of Submitters

The second type of data exclusion from the MPI process is the removal of specific years for submitters rather than the entire data set. This list was formed by analyzing the submitters across all years and determining how often multiple members were submitted under the same OrgID/CSUMID combination within each year. In these cases the majority of combinations were found to have been used over multiple individuals.

Every year CHIA works with submitters to obtain the best possible data. Several of these submitters improved their data quality and are included in the MPI process in more recent years.

ORG ID	EXCLUDED YEARS/VALUES
299	2011-2016
7249	2011-2018
7273	2015-2016
7473	2011-2015
10436	2011-2015
10632	2011-2013
10928	2011-2017
10933	2011-2018
10934	2013-2018
10935	2013-2014
11065	2013-2014, 2017-2018
11274	2012-2018
11426	2017-2018
11609	2014

3.   Record Level Exclusion of Submitters

The third type of data exclusion from the MPI process is specific records for submitters rather than the entire years.

This list was formed by analyzing the submitters across all years and determining how often multiple members were submitted under the same OrgID/CSUMID combination within a year. In these cases only a few combinations were found to have been used over multiple individuals.

ORG ID	EXCLUDED YEARS/VALUES
312	2 Instances of CSUMIDs have been removed from 201512
7268	1 Instance of CSUMID has been removed from 201512
10920	Remove specific IDs from source data
11237	6 Instances of CSUMIDs have been removed from 201712 and 201806