## CENTER FOR HEALTH INFORMATION AND ANALYSIS (CHIA)

# **CY2009-2013 INCURRED ALL-PAYER CLAIMS DATABASE (MA APCD) RELEASE 3.0 DOCUMENTATION GUIDE**

- Product File -

Issued: April 2015



Commonwealth of Massachusetts Center for Health Information and Analysis Áron Boros, Executive Director

# **Table of Contents**

# **Contents**

NTRODUCTION	∠
Section 1.0: History	5
1.1: Establishment of the Massachusetts APCD (MA APCD)	
1.2: MA APCD Release 3.0 Overview	
Section 2.0: MA APCD Data Collection Process	
2.1: Edits	
2.2: Variances	
2.3: Broad Caveats	
Section 3.0: Product File	
3.1: Types of Data Collected in the Product File:	
3.1.1: Product Identifiers	
3.1.2: Deductibles	
3.1.3: Dates	
3.2: Product Release File Structure:	
3.3: Product File Layout	
3.3.1 Restricted Release Elements	
3.3.2: Release Text File Column Titles	
3.3.3: File Layout Section Columns	
3.3.4: Product File Cleaning, Standardization, and Redaction.	

### INTRODUCTION

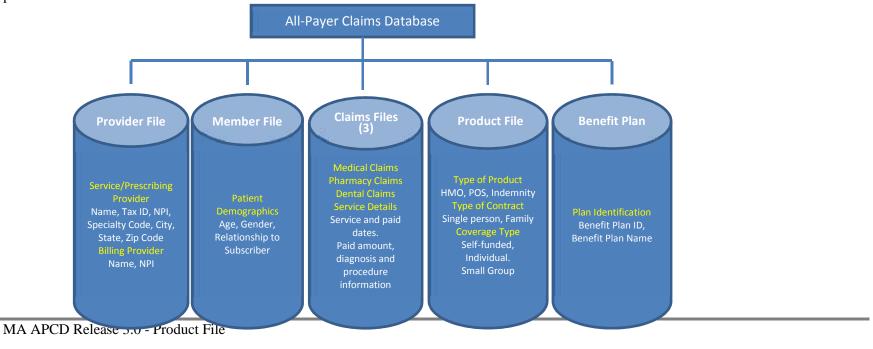
The Center for Health Information and Analysis (CHIA) was created to be the hub for high quality data and analysis for the systematic improvement of health care access and delivery in Massachusetts. Acting as the repository of health care data in Massachusetts, CHIA works to provide meaningful data and analysis for those seeking to improve health care quality, affordability, access, and outcomes.

To this end, the **All-Payer Claims Database** (**APCD**) contributes to a deeper understanding of the Massachusetts health care delivery system by providing access to accurate and detailed claims-level data essential to improving quality, reducing costs, and promoting transparency. This document is provided as a manual to accompany the release of data from the MA APCD.

The MA APCD is comprised of medical, pharmacy, and dental claims and information from the member eligibility, provider, product and benefit plan control files, that are collected from health insurance payers operating in the Commonwealth of Massachusetts. This information encompasses public and private payers as well as insured and self-insured plans.

APCD data collection and data release are governed by regulations which are available on the MA APCD website (see http://chiamass.gov/regulations/)

For ease of use, the Center for Health Information and Analysis (CHIA) has created separate documents for **each** APCD file type and one for the appendices—for a total of eight separate documents. All are available on the CHIA website.



## **Section 1.0: History**

## 1.1: Establishment of the Massachusetts APCD (MA APCD)

The first efforts to collect claim-level detail from payers in Massachusetts began in 2006 when the Massachusetts Health Care Quality and Cost Council (HCQCC) was established, pursuant to legislation in 2006, to monitor the Commonwealth's health care system and disseminate cost and quality information to consumers. Initially, data was collected by a third party under contact to the HCQCC. On July 1, 2009, the Division of Health Care Finance and Policy (DHCFP) assumed responsibility for receiving secure file transmissions, creating, maintaining and applying edit criteria, storing the edited data, and creating analytical public use files for the HCOCC. By July 2010, Regulations 114.5 CMR 21.00 and 114.5 CMR 22.00 became effective, establishing the APCD in Massachusetts.

Chapter 224 of the Acts of 2012, "An Act Improving the Quality of Health Care and Reducing Costs Through Increased Transparency, Efficiency and Innovation," created the Center for Health Information and Analysis (CHIA) which assumed many of the functions – including management of the MA APCD – that were previously performed by the Division of Health Care Finance and Policy (DHCFP).

According to Chapter 224, the purpose of the Massachusetts APCD is **Administrative Simplification**:

"The center shall collect, store and maintain such data in a payer and provider claims database. The center shall acquire, retain and oversee all information technology, infrastructure, hardware, components, servers and employees necessary to carry out this section. All other agencies, authorities, councils, boards and commissions of the commonwealth seeking health care data that is collected under this section shall, whenever feasible, utilize the data before requesting data directly from health care providers and payers. In order to ensure patient data confidentiality, the center shall not contract or transfer the operation of the database or its functions to a third-party entity, nonprofit organization or governmental entity; provided, however, that the center may enter into interagency services agreements for transfer and use of the data."

A Preliminary Release of the MA APCD – covering dates of service CY 2008-2010 and paid through February 28, 2011 – was first released in 2012. Release 3.0, to be available in early 2015, covers dates of service CY 2009-2013 (paid through June 30, 2014).

## 1.2: MA APCD Release 3.0 Overview

The MA APCD is comprised of data elements collected from all Private and Public Payers of eligible Health Care Claims for Massachusetts Residents. Data is collected in seven file types: Product (PR), Member Eligibility (ME), Medical Claims (MC), Dental Claims (DC), Pharmacy Claims (PC), Provider (PV), and Benefit Plan (BP) Control. Each is described separately in this user manual.

Highlights of the release include:

- Data is available for dates of service from January 1, 2009 to December 31, 2013 as paid through June 2014. Data submitted to CHIA after June 2014 is **NOT** included in the files.
- Release 3.0 contains more comprehensive and recently updated data, including resubmissions from several large carriers.
- Data elements are classified as either Level 2 or Level 3 data elements. Level 2 include data elements that pose a risk of re-identification of an individual patient. Level 3 data elements are generally either Direct Personal information, such as name, social security number, and date of birth, that uniquely identifies an individual or are among the 18 identifiers specified by HIPAA. Refer to the **File Layout** sections for listings of Level 2 and Level 3 data elements for each file.
- Public Use Files (PUFs), which are de-identified extracts of the Medical Claims (MC) and Pharmacy Claims (PC) files, will be release separately. The PUFs incorporate certain levels of aggregation and a much more limited list of elements to help ensure data privacy protection.
- Certain identifying or sensitive data elements are **Masked** in the release in order to protect personally identifiable information and allow for the linkage of data elements within the same file.
- Some data elements have been derived by CHIA from submission data elements or have been added to the database to aid in versioning and identifying claims (e.g. Unique Record IDs and status flags). Refer to the File Layout sections for detail

### Section 2.0: MA APCD Data Collection Process

The data collected from the payers for the MA APCD is processed by the **Data Compliance and Support** team. Data Compliance works with the payers to collect the data on a regular, predetermined, basis and ensure that the data is as complete and accurate as possible. The **Data Quality Assurance** and **Data** Standardization and Enhancement teams work to clean and standardize the data to the fullest extent possible. Data Standardization relies on external source codes (see Appendix 8) from outside government agencies, medical and dental associations, and other vendors to ensure that the data collectors properly utilized codes and lookup tables to make data uniform.

#### **2.1: Edits**

When payers submit their data to CHIA for the MA APCD, an Edits process is run on each file to check that the data complies with requirements for the file and for each data element in the file.

The automated edits perform an important data quality check on incoming submissions from payers. They identify whether or not the information is in the expected format (i.e. alpha vs. numeric), contains invalid characters (i.e. negative values, decimals, future dates) or is missing values (i.e. nulls). If these edits detect any issues with a file, they are identified on a report that is sent to the payer.

Data elements are grouped into four categories (A, B, C, and Z) which indicate their relative analytic value to CHIA and MA APCD users. Refer to the File **Layout** sections of each document to view the Edit Level for each Data Element:

- 'A' level fields must meet their MA APCD threshold percentage in order for a file to pass. There is an allowance for up to a 2% variance within the error margin percentage (depending on the data element). If any 'A' level field falls below this percentage it will result in a failed file submission for the payer and a discussion with their liaison regarding corrective action.
- The other categories (**B**, **C**, and **Z**) are also **monitored**, but the thresholds are not presently enforced.

More detailed APCD Version 3.0 File Edit documentation can be found at: http://chiamass.gov/apcd-data-submission-guides

### 2.2: Variances

The Variance process is a collaborative effort between the payer and CHIA to reach a mutually agreed upon threshold percentage for any data element which may not meet the MA APCD standard. Payers are allowed to request a lower threshold for specific fields, but they must provide a business reason (rationale) and, in some cases, a remediation plan for those elements. CHIA staff carefully reviews each request and follows up with a discussion with the payers about how to improve data quality, suggest alternative threshold rates or creating plans to reach threshold over time to improve reporting quality.

Once this process is complete, the variance template is loaded into production so that any submissions from the payer are held to the CHIA standard thresholds and any approved variances. The payer receives a report after each submission is processed which compares their data against the required threshold percentages. 'Failed' files are reviewed by CHIA liaisons and discussed with the payer for corrective action. (see Appendix 4)

### 2.3: Broad Caveats

Researchers using the MA APCD Release 3.0 data should be aware of the following:

- Due to the variance process, data quality may vary from one payer to another. (see Appendix 4)
- Claim Files submitted through June 2014 were accepted with relaxed edits. (Refer to the MA APCD Submission Guide for Edit information)
- The release files contain the data submitted to CHIA including valid and invalid values.
- Certain data elements were cleaned when necessary. Detail on the cleaning logic applied is described at the end of each file layout.
- Certain data elements were redacted to protect against disclosure of sensitive information.
- Some Release Data was manipulated to protect patient privacy:
  - O Assignment of linkage IDs to replace reported linkage identifiers (see Appendix 3).
  - Member Birth Year is reported as 999 for all records where the member age was reported as older than 89 years on the date of service.
  - Member Birth Year is reported as Null for all records where the member was reported as older than 115 years on the date of service.

### Section 3.0: Product File

As part of the MA APCD, payers are required to submit a **Product** file. Release 3.0 has one Product File that consists of aggregated and unduplicated records across multiple years.

Below are details on business rules, data definitions, and the potential uses of this data. For a full list of elements refer to the File Layout section.

### 3.1: Types of Data Collected in the Product File:

#### 3.1.1: Product Identifiers

CHIA has made a conscious decision to collect elementary identifiers that may be associated with a Product. The data in fields PR002 through PR008 can be used when analyzing Product data across payers. The identifiers will be used to help link Product data to the Member's Eligibility File.

#### 3.1.2: Deductibles

CHIA is collecting birth date information on each Subscriber and Member. This information is also useful with matching algorithms.

#### 3.1.3: Dates

CHIA collects two date fields for each Product record.

The Start and End Dates (PR009 and PR010) for each Product describes the dates the Product was active with the payer and usable by eligible members. For Products that were still active at the end of 2011, the End Date should be Null. For Products that were not active, but may still have claims being adjudicated against them, the End Date should be the End Date reported to the Division of Insurance OR the date the license was terminated.

# **3.2: Product Release File Structure:**

File Characteristic Topic	<del>Details</del> Clarificati	<u>on</u>				
Release File Format	sections. Empty or	Release files will be in an <b>asterisk delimited text file</b> in the same order as found in the File Layout sections. Empty or null data elements will have no spaces or characters between the asterisks. Each user will only receive the data elements requested and approved.				
Rows	products on separa	Each row is supposed to represent a unique instance of a Product. However, some payers have reported products on separate rows that differ only in aspects that are not specified in the Product file layout. Therefore, for some payers there may be appear to be duplicate rows, when in fact they are district product.				
Product Definition  Products or lines of business not inclu	PPO, Indemnity, e Product Line of Bu model. The data m File Lookup Table	tc. usiness Model (PR004) is collecteurs ust be submitted using a CHIA-p	d by a business model that it conforms to, e.g.: HMO, ed by the MA APCD to define the type of business provided lookup table, which is located in the Product e following:			
the lookup table for PR004	Element	Element Name	Submission Guideline			
	PR004	Product Line of Business	ZZ			
	PR007	Other Product benefit Description	Payers may enter the name of the business			
	By reporting the Model Code of ZZ (mutually defined by CHIA and the Payers) the Payer will be able to report the name of the business model in PR007. CHIA realizes that payers store their Product data in a variety of formats and data structures. CHIA feels this methodology will provide the most flexibility to analyze Product data.					

### 3.3: Product File Layout

#### 3.3.1 Restricted Release Elements

- Each **row** in the release file contains one record of the indicated file type. There is an **asterisk-delimited field** in each row for every data element listed in the Restricted Release sections for each file type.
- Data Elements will be delimited in the order displayed in the File Layout sections of this document.
- **Empty** or **null** data elements will have no spaces or characters between the asterisks.
- Lookup Tables: Element-specific Lookup Tables are included in this document after each File Type Layout section.
- A Carrier-Specific Master Lookup table is included with each data extract. Refer to the Carrier-Specific Reference and Linking sections in this document for more information.
- **External Code Sources** are listed in Appendix 9.
- Masked Elements: For the Data Release, some of the data elements have been Masked to provide confidentiality for Payers and Providers, and individuals, while allowing for linkage between claims, files, and lookup tables. Refer to the Data Protection/Confidentiality and Linkage sections of the Appendices for more information.

#### 3.3.2: Release Text File Column Titles

Appendix 610: Release File Column Names included in this document lists the column name for each data element in the Level 2 and Level 3 release files. The text files exported from the APCD SQL Database include these SQL column names in the first row.

#### 3.3.3: File Layout Section Columns

- Data Element: The code name of the element, with reference to the Regulation and the Submission files received by CHIA from Payers. The first two digits refer to the File Type and the following numbers to the ordering in the Submission Files.
- Data Element Name: Name of the element.
- Format/Length: Maximum Length of the data column in the APCD's SQL Server database at CHIA.
- **Description:** Description of the element; **additionally** the lookup table is included where applicable.
- Additional Element Description: Additional information about the element in the release.
- Edit Level: Level of enforcement of the data element's requirements by CHIA on Payer Submissions. Refer to the Edits section of this document.
- %: The expected percentage of validity for instances of the element in each submission file by the Payer.

			MA APCD Produ	ct File – Level 2 Data Eleme	ents		
Data Element	Data Element Name	Format / Length	Description	Element Submission Guideline	Additional Element Description	%	Edit Level
Derived- PR1	Release ID	int-NULL		Unique record ID derived specifically for this release file type	With each release file type table this number is reset to 1 and sequentially incremented by one for every record released	N/A	N/A
Derived- PR2	Submission Year and Month	int-NULL				N/A	N/A
Derived- PR3	Medicaid Indicator	int-NULL				N/A	N/A
PR001	Product ID number	varchar[30]	Product Identification	Report the submitter- assigned identifier that uniquely defines this product. This identifier is used to link this Product line with its attributes to eligibility and claim lines.		100%	A0
PR003	Carrier License Type	char[3]	Carrier License Type  BLU Blue Cross and Blue Shield Licensee COM Commercial Carrier HMO Health Maintenance Organization MAO Medicare Advantage Organization PBM Pharmacy Benefit Manager SCO Senior Care Option TPA Third Party Administrator	Report the code that defines the License Type of this Product.  EXAMPLE: COM = Commercial	A code that defines the license type associated with the Product filing with the Massachusetts Division of Insurance.	100%	A0

			MA APCD Produ	ct File – Level 2 Data Elem	ents		
Data Element	Data Element Name	Format / Length	Description	Element Submission Guideline	Additional Element Description	0/0	Edit Level
			176 Chapter 176 OTH Other License Ty				
PR004	Product Line of Business Model	char[2]	Line of Business / Insurance Model the Product relates to:  12 Preferred Provider Organization (PPO) 13 Point of Service (POS) 14 Exclusive Provider Organization (EPO) 15 Indemnity Insurance 16 Health Maintenance Organization (HMO) Medicare Advantage 20 Medicare Advantage PPO 21 Medicare Advantage Private Fee for Service AC Accident Only BH Basic Hospital CH CHAMPUS DM Dental Maintenance Organization DS Disability HC HMO - Closed HO HMO - Open IN Individual LM Liability Medical MC Medicaid FFS MO Medicaid Managed Care Organization	Report the code that defines the Line of Business model that this product follows.  EXAMPLE: 12 = PPO	A code that defines a product's business model as defined by the carrier or its designee. Value of ZZ (Other) should correspond to non-insurance vendors; Claim Reprocessors or Re-pricers, Computer Leasing, etc.	100%	A0

	MA APCD Product File – Level 2 Data Elements								
Data Element	Data Element Name	Element Length Guideline		%	Edit Level				
			MP Medicare Primary MR Medicare MS Medicare Secondary Plan OF Other Federal Program (e.g. Black Lung) PC Medicaid Primary Care Clinician Plan PR Preferred Provider Organization (PPO) QM Qualified Medicare Beneficiary/SLMB SA Self-Administered Group SC Senior Care Option SP Supplemental Policy TF HSN Trust Fund TV Title V UN Unemployment VA Veterans Administration Plan VS Vision WC Workers' Compensation ZZ Other Code 12 Preferred Provider Organization (PPO) 13 Point of Service (POS) 14 Exclusive Provider Organization (EPO) 15 Indemnity Insurance 16 Health Maintenance Organization (HMO)						

	MA APCD Product File – Level 2 Data Elements									
Data Element	Data Element Name	Element Length Guideline		%	Edit Level					
			Medicare Advantage PPO 21 Medicare Advantage Private Fee for Service AC Accident Only BH Basic Hospital CH CHAMPUS DM Dental Maintenance Organization DS Disability HC HMO - Closed HO HMO - Open IN Individual LM Liability Medical MC Medicaid FFS MO Medicaid Managed Care Organization MP Medicare Primary MR Medicare MS Medicare Secondary Plan OF Other Federal Program (e.g. Black Lung) PC Medicaid Primary Care Clinician Plan PR Preferred Provider Organization (PPO) QM Qualified Medicare Beneficiary/SLMB SA Self-Administered Group SC Senior Care Option SP Supplemental Policy							

			MA APCD Produ	ıct File – Level 2 Data Eler	nents		
Data Element	Data Element Name	Format / Length	Description	Element Submission Guideline	Additional Element Description	0/0	Edit Level
DD005		shouf 41	TF HSN Trust Fund TV Title V UN Unemployment VA Veterans Administration Plan VS Vision WC Workers' Compensation ZZ Other	Deposit the ends that	A so de that de Cines a gree de et?	1000/	40
PR005	Insurance Plan Market	char[4]	Insurance Plan Market Code  GPOS Group - POS GCOB Group COBRA GCCH Group- Commonwealth Choice GEMP Group-Employer GFED Group-Federal GGIC Group-GIC GMMK Group-Merged Market GMUN Group- Municipality GPRT Group-Retiree GSC0 Group-Senior Care Option GUNN Group-Union HEXC Health Exchange ICCA Individual - Commonwealth Care ICCH Individual Closed ICOB Individual COBRA	Report the code that defines the market this product is sold into.  EXAMPLE: GEMP = Group - Employer	A code that defines a product's business model as defined by the carrier or its designee. Value of ZZ (Other) should correspond to non-insurance vendors; Claim Reprocessors or Re-pricers; Computer Leasing; etc.	100%	A0

	MA APCD Product File – Level 2 Data Elements								
Data Element	Data Element Name	Format / Length	Description	Element Submission Guideline	Additional Element Description	%	Edit Level		
			ISCO Individual Senior Care Option IYGA Individual Young Adult MCRA Medicare Part A MCRB Medicare Part B MCRC Medicare Part C MCRD Medicare Part D MEDX MediGap/Medicare Supplemental/Medex ITHR Other OTMC Other Medicare STUD Student COBR COBRA GRUP Group						
PR006	Product Benefit Type	int[1]	Benefit Options  1 Medical Only 2 Pharmacy Only 3 Medical and Pharmacy bundled 4 Dental 5 Behavioral Health 6 Vision 7 Accident Only 8 Medical Comprehensive 0 Other	Report the value that defines the types of benefits covered under this product. <b>EXAMPLE:</b> 1 = Medical Only	Numeric indicator that reports a benefit selection or a product-range offering as defined by the carrier or its designee.	100%	A0		
PR008	Risk Type	int[1]	Risk Options  1 Fully Insured 2 Self-Insured 3 Product available to risk and self-insured accounts	Report the value that best describes the risk model that defines how eligibilities are insured under this product line. <b>EXAMPLE:</b> 1 = Fully	Numeric indicator that reports the product development attribute that defines a risk assignment.	100%	A0		

			MA APCD Produ	ct File – Level 2 Data Eleme	nts		
Data Element	Data Element Name	Format / Length	Description	Element Submission Guideline	Additional Element Description	%	Edit Level
			0 Other	Insured			
PR009	Product Start Date	datetime- NULL	Product Start Date	Report the first date that this product is active in CCYYMMDD Format.	First date that a product is eligible for Member enrollment.	100%	A0
PR010	Product End Date	int[8]	Product End Date	Report the last date that this product is active in CCYYMMDD Format. If product is still active do not report any value here.	(YYYY-MM-DD 00:00:00.000)	100%	В
PR011	Product Active Flag	int[1]	Indicator - Active Product  1 Yes 2 No 3 Unknown 4 Other 5 Not Applicable	Report the value that defines the element. <b>EXAMPLE:</b> 1 = Yes, the Product is active	Last date that product is active for Member enrollment. (YYYY-MM-DD 00:00:00.000)	100%	A2
PR012	Annual Per Person Deductible Code	char[3]	Per Person Deductible bandwidth reporting  000 No per person deductible 001 Deductible Total under \$1,000 002 Deductible Total of \$1,000 thru \$1,999 003 Deductible Total of \$2,000 thru \$2,999 004 Deductible Total greater than \$3000 999 Not Applicable	Report the value that defines the Total Per Person Deductible for all benefits under this product. <b>EXAMPLE:</b> 000 = Plans with no Per Person Deductible	Numeric indicator that reports active vs. inactive products for the date span indicated in Product Start and End Dates.	100%	A2

			MA APCD Produ	ct File – Level 2 Data Eleme	nts		
Data Element	Data Element Name	Format / Length	Description	Element Submission Guideline	Additional Element Description	%	Edit Level
PR013	Annual Per Family Deductible Code	char[3]	char[3] Per Family Deductible bandwidth reporting  000 No per family deductible 001 Deductible Total under \$1,000 002 Deductible Total of \$1,000 thru \$1,999 003 Deductible Total of \$2,000 thru \$2,999 004 Deductible Total of \$3,000 thru \$3,999 005 Deductible Total of \$4,000 thru \$4,999 006 Deductible Total of \$5,000 thru \$5,999 007 Deductible Total greater than \$6,000 999 Not Applicable	Report the value that defines the Total Per Family Deductible for all benefits under this product.  EXAMPLE: 000 = Plans with no Per Family Deductible	Value that represents the Total Per Person Deductible for all benefits under this product for the date span indicated in Product Start and End Dates.	100%	A2
PR014	Coordinated Care model	int[1]	Indicator - Clinical Coordination  1 Yes 2 No 3 Unknown 4 Other 5 Not Applicable	Report the value that defines the element. <b>EXAMPLE:</b> 1 = Yes,  Member's care is clinically coordinated/managed.	Value that represents the Total Per Family Deductible for all benefits under this product for the date span indicated in Product Start and End Dates	100%	A2
PR017	NAIC Code	int[5]	National Association of Insurance Commissioners' Code			100%	A2

			MA APCD Produ	ict File – Level 3 Data Eleme	ents		
Data Element	Data Element Name	Format / Length	Description	Element Submission Guideline	Additional Element Description	%	Edit Level
PR002	Product Name	varchar[70]	Submitter defined Product name	Report a unique name for every Product in a Carrier's system. For Products with identical names, it is required that the Submitter add a refining 'element' to create unique Product Names that align to unique Product ID Numbers. This refining element can be numeric, alpha or alpha-numeric.	Unique name assigned to Product by Carrier/Submitter	100%	С
PR007	Other Product Benefit Description	varchar[80]	Benefit Description	Report the Other Product description when the product's Product Benefit does not fall within the standard listing for PR006 Product Benefit Type <b>EXAMPLE:</b> Chiropractic	Refining description applied by carrier or its designee when PR006 was reported as 0 (Other).	100%	A2
PR015	Other Product Line of Business Model	varchar[30]	Defines Product Line of Business when Other is selected	Report the Other Product Line of Business Model here when PR004 reports ZZ.		98%	A2
PR016	Other Risk Type	varchar[30]	Defines Risk Type Other is selected	Report the Other Risk Type here when PR008 reports 0 <zero>.</zero>		98%	A2
PR899	Record Type	char[2]	File Type Identifier		All	100%	A0

# 3.3.4: Product File Cleaning, Standardization, and Redaction

MA APCD Product File Cleaning Logic, by Element								
Element	Element Data Element Name Format/Length Description Cleaning Logic							
N/A								

MA APCD Product Claims File SSN Redaction, by Element					
Element	Data Element Name	Format/Length	Description		
PR015	Other Product Line of Business Model	Varchar[30]	Other Product Line of Business Model		
PR016	Other Risk Type	Varchar[30]	Other Risk Type		

MA APCD Product File Reidentification, by Element				
Element	Data Element Name	Format/Length	Description	
PR001	Product ID number	varchar[6]	Product ID number (linking)	



## **Center for Health Information and Analysis**

501 Boylston Street, Boston MA 02116 617-701-8100 www.chiamass.gov