

Chronic Condition Data Warehouse

Receiving CCW Data via External Media

Contents

Change Log.....	iii
I. Overview.....	1
II. Hard Drive Contents	1
A. The Extract Process.....	1
B. Decompression and Decryption.....	1
1. If you are using a Windows operating system	2
2. If you are using a non-Windows operating system	2
C. Other Documents Included on Your Delivery Media.....	3
III. Creating Usable Data Files	3
A. Joining Summary File Segments.....	3
B. Joining Summary Files and Crosswalk Files.....	4
C. Joining Summary Files and Part D Event Files.....	4
D. Joining Summary Files and Base Claim Records	5
E. Joining Institutional Base Claims and Revenue Center Records.....	6
F. Joining Institutional Base Claims and Reference Code Files.....	7
G. Joining Non-Institutional Base Claims and Line Item Records.....	7
IV. Where to Get More Help	9

**Chronic Condition Data Warehouse
Receiving CCW Data via External Media**

Change Log

Date	Changed By	Changes	Version
11/07/2012	K. Hutchins	Created document	1.0

I. Overview

The purpose of this guide is to provide step-by-step instructions for loading your Chronic Condition Data Warehouse (CCW) data extracts onto the hardware of your choice, opening and decrypting the files, and joining the files into usable data sets.

II. Hard Drive Contents

A. The Extract Process

When your data use agreement (DUA) has been approved, a data analyst at Buccaneer prepares a query to extract the data specified in your DUA from the CCW. The data will be written to a series of flat data files, with certain sensitive fields encrypted according to the restrictions specific to your DUA. Each separate data file is packaged with three additional files needed to use the data, and each resulting group of four files is encrypted and compressed into a Self-Decrypting Archive (SDA) executable file. These SDAs are copied to an external hard drive or optical media (DVD/CD) and shipped to you.

A minimal set of instructions for installing the data from the shipping media, including the password needed to open the SDAs, will be sent to you separately by email. If you cannot locate this email, or believe you have never received it, please contact the CCW Analytic Team via email (cmsdata@gdit.com) or telephone (866-766-1915) to have the material retransmitted.

B. Decompression and Decryption

Your extract package contains a Windows-compatible external USB hard drive or DVD/CD onto which a number of executable files (SDAs) have been copied. Please follow these specific instructions to decrypt and decompress these files and transfer them to the hardware destination where you will use them for analysis.

Chronic Condition Data Warehouse Receiving CCW Data via External Media

1. If you are using a Windows operating system

- a) *If you have received your data on external hard drive, attach the hard drive to your Windows computer via the included USB cable. Verify that your computer has recognized the hard drive via Start>Computer (Win7) or the equivalent commands in earlier Windows versions. It should appear as a Hard Drive with an assigned drive letter.*
- b) *If you have received a DVD/CD containing your data, insert the disk into the optical drive of your Windows computer.*
- c) *(USB drive recipients only) Verify that there is enough space remaining on the USB hard drive to decrypt and decompress the data. The information on the amount of disk space necessary for decryption and decompression is included in the cover letter that came with your shipment.*
- d) *If there is sufficient space, proceed to Step (e). Otherwise, or if you have received your data on DVD/CD, copy the SDAs to the target Windows platform where you intend to store the decompressed files.*
- e) *Obtain the SDA password from the email sent by the GDIT HI Specialist.*
- f) *Run each executable file by double-clicking the file name or right-clicking and choosing "Run" from the context menu. Enter the password when prompted.*

2. If you are using a non-Windows operating system

- a) *Attach the hard drive to your system via the included USB cable. Verify that your system has recognized the hard drive.*
- b) *Verify that there is enough space remaining on target disk to decrypt and decompress the data. The information on the amount of disk space necessary for decryption and decompression is included in the cover letter that came with your shipment.*
- c) *Upload the SDAs to the target platform.*
- d) *Obtain the SDA password from the email sent by the GDIT HI Specialist.*
- e) *Run each executable file. Enter the password when prompted.*

:

Chronic Condition Data Warehouse Receiving CCW Data via External Media

For each executable file, you will produce four files, named according to the following convention:

- `datarecordtype` is the type of claim, summary or other data that was extracted; e.g., “outpatient_base_claims” or “snf_revenue_center”
- `resnnnnn` is the DUA under which your data was released
- `reqnnnnn` is the CCW request number under which your data was processed
- `yyyy` is the year of data included in the file. If you have requested multiple years of the same type of claim, each year will be in a separate file

<code>datarecordtype_res12345_req0678_yyyy.exe</code>	Original SDA executable file
<code>datarecordtype_res12345_req0678_yyyy.dat</code>	Flat data file
<code>datarecordtype_res12345_req0678_yyyy.fts</code>	File Transfer Summary file
<code>datarecordtype_read_v6.sas</code>	SAS input statements necessary to read the flat data file into a V6 SAS data set
<code>datarecordtype_read_v8.sas</code>	SAS input statements necessary to read the flat data file into a V8 SAS data set

C. Other Documents Included on Your Delivery Media

A directory named “File Record Layouts” contains data dictionaries of the file types included in your request. These files are typically either PDFs or Excel workbooks.

III. Creating Usable Data Files

The SAS input statements necessary to convert your flat data files are automatically extracted when you execute each SDA. Two versions are provided: those files ending with `_read_v6` are compatible with earlier SAS versions where variable names were limited to eight bytes, while the `_read_v8` files use the 32 byte limit introduced in later versions of SAS. Once you have created your permanent SAS data sets and placed them in the library or libraries where you wish to store them, you will have to perform some or all of the following steps to make your data usable, depending upon what kind of claims, summary files, and/or crosswalks you are using.

A. Joining Summary File Segments

Summary files contain one record per beneficiary. The unique identifier variable is called `BENE_ID`. This variable is encoded before your data is sent to you, with a key unique to your DUA. Data released under your DUA can be merged by `BENE_ID`, but you cannot merge this data with CCW

Chronic Condition Data Warehouse Receiving CCW Data via External Media

data released under any other DUA, because that data has been encoded using a different unique key.

The Master Beneficiary Summary File consists of separate segments, including the Base File (A/B or A/B/D), Chronic Condition, Cost and Use, and National Death Index. (Your data request may not have included all these segments). These segments can all be merged together on BENE_ID. This is a one-to-one matching -- each BENE_ID is represented in the file only once.

(The BENE_ID and other ID variable values listed in the examples in this and the following sections are fictitious. They are meant as illustration only.)

BENE_ID	Master Bene Summary Base	Chronic Condition Segment	Cost and Use Segment
1472	→ 1472	→ 1472	→ 1472
5784	→ 5784	→ 5784	→ 5784
9837	→ 9837	→ 9837	→ 9837
10926	→ 10926	→ 10926	→ 10926
86745	→ 86745	→ 86745	→ 86745
723401	→ 723401	→ 723401	→ 723401

B. Joining Summary Files and Crosswalk Files

Your approved DUA may have included permission to merge CCW data with data already in your possession that contains a different index variable, such as Health Insurance Claim number (HIC), Social Security number (SSN), Medicare Current Beneficiary Summary BASE_ID, Beneficiary Identification number (BID) or assessment resident identification number (RES_ID). If that is the case, your data release includes a crosswalk file or files that will enable you to join your CCW data to the other data sources. Simply merge the crosswalk file to your other summary files on BENE_ID, and to your existing data on the appropriate index variable.

CCW Extract Data BENE_ID	CCW Crosswalk File BENE_ID	SSN	Outside Data SSN
1472	→ 1472	123456789	→ 123456789
5784	→ 5784	234567890	→ 234567890
9837	→ 9837	345678901	→ 345678901
10926	→ 10926	456789012	→ 456789012
86745	→ 86745	567890123	→ 567890123
723401	→ 723401	678901234	→ 678901234

C. Joining Summary Files and Part D Event Files

A record in a Part D Event file corresponds to a single fill of a prescription, making them similar to base claims records in Part A and Part B data. There are potentially many Part D Event records for a single beneficiary, resulting in a one-to-many match when these files are joined.

Chronic Condition Data Warehouse Receiving CCW Data via External Media

CCW Summary File		CCW Part D Event	
BENE_ID		BENE_ID	PDE_ID
1472	→	1472	8567346
	→	1472	2458374859
	→	1472	43756432
5784	→	5784	9384756112
	→	5784	4736475849
9837	→	9837	547362
10926	→	10926	85948302938
	→	10926	847382732
	→	10926	473625364
	→	10926	9485746
86745	→	86745	5049384594
	→	86745	584738263
	→	86745	485968463745
	→	86745	732647384
723401	→	723401	37463856
	→	723401	192837845

D. Joining Summary Files and Base Claim Records

Both institutional (Part A) and non-institutional (Part B) claims come in multiple files per claim type. At the highest level of aggregation, both Part A claims (Inpatient, Skilled Nursing Facility, Home Health Agency, Hospice, and Outpatient Hospital) and Part B claims (Carrier and Durable Medical Equipment) contain a base claim record. There are potentially multiple claims per beneficiary, so when summary files are joined to base claim files, a one-to-many merge results.

Base claims records contain the unique index variable CLM_ID.

CCW Summary File		CCW Base Claim	
BENE_ID		BENE_ID	CLM_ID
1472	→	1472	903475639485723
	→	1472	584730495817345
	→	1472	675849304958674
5784	→	5784	192039485746372
	→	5784	102938102938475
9837	→	9837	483729102938475
10926	→	10926	384928394059483
	→	10926	837263748273764
	→	10926	987645372615243
	→	10926	672938473264537
86745	→	86745	116273628374659
	→	86745	664738447564732
	→	86745	827364837465743

Chronic Condition Data Warehouse Receiving CCW Data via External Media

	→	86745	867593847584756
723401	→	723401	172637485938394
	→	723401	784736283948571

E. Joining Institutional Base Claims and Revenue Center Records

Revenue center records supply more detailed information about Part A claims. Potentially, multiple revenue center records exist for a single base claim, resulting in a one-to-many match when these two file types are joined. A record in the revenue center file is uniquely identified by the combination of CLM_ID and CLM_LINE_NUM (CLM_LN in SAS V6).

CCW Summary File		CCW Base Claim			CCW Revenue Center
BENE_ID		BENE_ID	CLM_ID		CLM_LN (v6) CLM_LINE_NUM (v8)
1472	→	1472	903475639485723	→	1
	→	1472	584730495817345	→	1
			584730495817345	→	2
1472	→	1472	675849304958674	→	1
			675849304958674	→	2
			675849304958674	→	3
5784	→	5784	192039485746372	→	1
			192039485746372	→	2
5784	→	5784	102938102938475	→	1
9837	→	9837	483729102938475	→	1
			483729102938475	→	2
10926	→	10926	384928394059483	→	1
			384928394059483	→	2
			384928394059483	→	3
10926	→	10926	837263748273764	→	1
			837263748273764	→	2
10926	→	10926	987645372615243	→	1
	→	10926	672938473264537	→	1
10926			672938473264537	→	2
86745	→	86745	116273628374659	→	1
	→	86745	664738447564732	→	1
			664738447564732	→	2
86745			664738447564732	→	3
	→	86745	827364837465743	→	1
			827364837465743	→	2
86745	→	86745	867593847584756	→	1
			867593847584756	→	2
723401	→	723401	172637485938394	→	1
	→	723401	784736283948571	→	1
			784736283948571	→	2

F. Joining Institutional Base Claims and Reference Code Files

Reference code files contain information regarding special conditions which may affect payer processing. There are four types of reference code files: condition, occurrence, span, and value code. Reference code files contain one observation for each base claim, so when base claims are joined to reference code files it results in a one-to-one match. The unique ID variable in all files is CLM_ID.

Summary File	CCW Base Claim		Condition	Occurrence	Span	Value
BENE_ID	BENE_ID	CLM_ID	CLM_ID	CLM_ID	CLM_ID	CLM_ID
1472	→	1472	903475639485723	903475639485723	903475639485723	903475639485723
		1472	584730495817345	584730495817345	584730495817345	584730495817345
		1472	675849304958674	675849304958674	675849304958674	675849304958674
5784	→	5784	192039485746372	192039485746372	192039485746372	192039485746372
		5784	102938102938475	102938102938475	102938102938475	102938102938475
9837	→	9837	483729102938475	483729102938475	483729102938475	483729102938475
10926	→	10926	384928394059483	384928394059483	384928394059483	384928394059483
		10926	837263748273764	837263748273764	837263748273764	837263748273764
		10926	987645372615243	987645372615243	987645372615243	987645372615243
		10926	672938473264537	672938473264537	672938473264537	672938473264537
86745	→	86745	116273628374659	116273628374659	116273628374659	116273628374659
		86745	664738447564732	664738447564732	664738447564732	664738447564732
		86745	827364837465743	827364837465743	827364837465743	827364837465743
		86745	867593847584756	867593847584756	867593847584756	867593847584756
723401	→	723401	172637485938394	172637485938394	172637485938394	172637485938394
		723401	784736283948571	784736283948571	784736283948571	784736283948571

G. Joining Non-Institutional Base Claims and Line Item Records

Line item records supply more detailed information about Part B claims. Potentially, multiple line item records exist for a single base claim, resulting in a one-to-many match when these two file types are joined. A record in the line item file is uniquely identified by the combination of CLM_ID and LINE_NUM.

Chronic Condition Data Warehouse Receiving CCW Data via External Media

CCW Summary File BENE_ID		CCW Base Claim			CCW Line Item LINE_NUM (v8)
		BENE_ID	CLM_ID		
1472	→	1472	903475639485723	→	1
	→	1472	584730495817345	→	1
			584730495817345	→	2
	→	1472	675849304958674	→	1
			675849304958674	→	2
			675849304958674	→	3
5784	→	5784	192039485746372	→	1
			192039485746372	→	2
	→	5784	102938102938475	→	1
9837	→	9837	483729102938475	→	1
			483729102938475	→	2
10926	→	10926	384928394059483	→	1
			384928394059483	→	2
			384928394059483	→	3
	→	10926	837263748273764	→	1
			837263748273764	→	2
	→	10926	987645372615243	→	1
	→	10926	672938473264537	→	1
			672938473264537	→	2
86745	→	86745	116273628374659	→	1
	→	86745	664738447564732	→	1
			664738447564732	→	2
			664738447564732	→	3
	→	86745	827364837465743	→	1
			827364837465743	→	2
	→	86745	867593847584756	→	1
			867593847584756	→	2
723401	→	723401	172637485938394	→	1
	→	723401	784736283948571	→	1
			784736283948571	→	2

Additional information about manipulating CCW data files to create SAS analytic datasets is available on the CCW website at: <http://www.ccwdata.org/>

Please consult **CCW Technical Guidance: Getting Started with CMS Administrative Research Files** for more detailed information on types of claims files, data anomalies, and samples of SAS code.

IV. Where to Get More Help

Additional information about CCW extract data variables	Research Data Assistance Center website http://www.resdac.org
<ul style="list-style-type: none">• Technical Guidance documents• User Guides• Data Dictionaries• Analytic FAQs	Chronic Condition Data Warehouse website http://www.ccwdata.org/

If you are unable to complete any of the instructions in this guide please email cmsdata@gdit.com or call (866) 766-1915 and select Option 3. Please identify yourself as a data extract user who received data delivery on external media, so that your help request can be routed to the person best able to assist you.