



Submission of Medicare Data Finder and Crosswalk Files

September 2011

Finder File Encryption Policy

This section describes the minimum standards of data security that must be met by all finder files sent to Buccaneer.

Encryption Requirements

To maintain and assure the security of identifiable information shipped to Buccaneer, all finder files containing secure information must be encrypted with an Advanced Encryption Standard (AES) of at least 256-bit encryption algorithm. This is a block cipher adopted as an encryption standard by the U.S. government. The National Security Agency (NSA) has deemed the use of 256-bit AES encryption algorithm as secure enough for U.S. government top secret information.

Software Options

Files should be encrypted using software programs capable of 256-bit encryption. Buccaneer is unable to install all encryption tools in use; therefore, the finder file should be:

- A self-decrypting archive (SDA) or the file is WinZIP (version < 12.0) or PKZip compatible
- Decrypted on Solaris or Windows using a password only. (Buccaneer does not have other platforms available)

Password and Shipping Requirements

The files must be password protected with a password no less than 10 digits in length and must contain a combination of letters and numbers. The data shipment must be sent by a courier service with tracking capabilities, such as FedEx, UPS, DHL, or Registered Mail options. Do not include the password along with the data shipment, but instead, email the password to Buccaneer at CMSdata@vangent.com.

Ship Encrypted Finder Files to:

Cindy Weigel
Buccaneer, A Vangent Company
Software Development and Informatics
1401 50th St., Suite 200
West Des Moines, IA 50266



Finder Files and Crosswalks

The Chronic Condition Data Warehouse (CCW) can extract data from Medicare’s Research Identifiable Files (RIFs) for individual beneficiaries using information provided through finder files submitted along with data requests approved by CMS. Finder files may consist of the following types:

- *BENE_IDs or BIDs* - Beneficiary IDs received from a previous data shipment*
- *HICs* - Health Insurance Claim numbers
- *SSNs* - Social Security numbers
- *RES_ID / State Code* - Identifies resident in the national repository
- *UPINs* - Unique Physician Identification Number

*Data delivered by Buccaneer contains *BENE_ID*; whereas data delivered prior to January 2009 by Acumen LLC contains *BID*.

The following describes the required format of these files, along with the rules applied to extract the data. In addition, this document specifies the form of “crosswalk” (or “conversion” files) that link the *submitted* beneficiary ID with the *new* BENE_IDs assigned in the new shipment of Medicare data.

BENE_ID/BID Finder Files

BENE_IDs (or BIDs in data delivered prior to January 2009 by Acumen LLC) are a new way of identifying the claims associated with particular beneficiaries in deliveries of RIF data. The CCW delivers RIFs with a BENE_ID data element incorporated as a part of each claim. A BENE_ID designates a specific Medicare beneficiary, with the same BENE_ID identifying this beneficiary in all claims included in the shipment. So, claims for an individual can readily be linked using the person’s BENE_ID, across all categories of claims and all years making up the shipment. The assignment of BENE_IDs is unique for each Data Use Agreement (DUA).

Identifying a beneficiary’s claims from previous shipments requires submission of a BENE_ID finder file. A finder file identifying beneficiaries in a previous job must include only one informational item per individual which would be the beneficiary’s BENE_ID. The person’s gender and date of birth are not necessary for a BENE_ID finder file since each BENE_ID has already been uniquely assigned.

A BENE_ID finder file must be in CSV (ASCII comma delimited) format structured as follows:

1st line

List the variable name as follows:

BENE_ID



2nd line and beyond

Data for BENE_ID for each beneficiary:

BENE_ID: 15 digit CHAR identifying beneficiary

A BID finder file must be in CSV (ASCII comma delimited) format structured as follows:

1st line

List the variable name as follows:

BID_ 'Control Number' _ 'Phase Number'

2nd line and beyond

Data for BID for each beneficiary:

BID: 9 digit number identifying beneficiary

A carriage return must end each line.

BENE_ID to BID Crosswalk Files

In January 2009, the RDDC transitioned from Acumen LLC to Buccaneer. RIF data delivered prior to January 2009 by Acumen LLC contained a beneficiary identifier labeled BID. Data delivered by Buccaneer contains a beneficiary identifier labeled BENE_ID. Researchers interested in linking data delivered by Acumen LLC with data delivered by Buccaneer must request a BENE_ID to BID crosswalk on the RIF Specifications Worksheet.

The DUA number and job Control Number under which the BIDs were processed should be included on the RIF Specification Worksheet.

	Long Name	Short Name	Format	Start Column	Length
1	CONTROL_NUMBER	CNTRLNUM	NUM	1	4
2	BID	BID	CHAR	5	9
3	BENE_ID	BENE_ID	CHAR	14	15

HIC Finder and Crosswalk Files

Submitting a HIC finder file requires a listing of beneficiaries' HICs. Buccaneer will extract data for all Medicare beneficiaries identified by the submitted HICs. The requestor may also include individual's gender and DOB information along with HICs in the finder file. While submitting these additional variables does not change the process of strictly matching to HIC, the crosswalk returned to the requestor will show the matches made to the gender and DOB variables.



The requestor does have the option of restricting the linkage of HIC finder file beneficiaries to include only those who strictly match on their gender and/or DOB as well as on HIC. If the requestor desires to have these additional selection criteria imposed, this must be clearly stated on the RIF Specification Worksheet.

A HIC finder file must be in CSV (ASCII comma delimited) format structured as follows:

1st line

List of variable names as follows: HIC, Gender, DOB

2nd line and beyond

Data for above variables for each record:

HIC: 12 character alpha-numeric code identifying beneficiary
Gender: 'M' or '1' = male, 'F' or '2' = female and '0' = unknown
DOB: mm/dd/yyyy
(Gender and DOB are optional variables)

A carriage return must end each line.

When submitting HICs, the following guidelines will reduce the risk of misinterpretation of the finder file:

- 10 or 11 length (CMS HIC or RRB Number)
- Letters, Numbers, or leading '{' Only.
- Left Justified

The HIC crosswalk file will be a fixed length text format structured as follows:

1st line

List of variable names as follows:

HIC, BENE_ID

2nd line and beyond

Data for above variables for each beneficiary

The second variable in this file, BENE_ID, refers to the BENE_ID assigned to a beneficiary in the new shipment.



SSN Finder and Crosswalk Files

Submitting a SSN finder file requires a listing of beneficiaries' SSNs. Buccaneer will extract data for all Medicare beneficiaries identified by the submitted SSNs. The requestor may also include individual's gender and DOB information along with SSNs in the finder file. While submitting these additional variables does not change the process of strictly matching to SSN, the crosswalk returned to the requestor will show the matches made to the gender and DOB variables.

The requestor does have the option of restricting the linkage of SSN finder file beneficiaries to include only those who strictly match on their gender and/or DOB as well as on SSN. If the requestor desires to have these additional selection criteria imposed, this must be clearly stated on the RIF Specification Worksheet.

A SSN finder file must be in CSV (ASCII comma delimited) format structured as follows:

1st line

List of variable names as follows: SSN, Gender, DOB

2nd line and beyond

Data for above variables for each beneficiary:

SSN: 9 digit number identifying beneficiary
Gender: 'M' or '1' = male, 'F' or '2' = female and '0' = unknown
DOB: mm/dd/yyyy
(Gender and DOB are optional variables)

A carriage return must end each line

When submitting SSNs, the following guidelines will reduce the risk of misinterpretation of the finder file:

- Length should be 9 and exclude all dashes
- Numbers Only
- Alphanumeric Formatting (Leading Zeroes Preserved)
- Left Justified

The SSN crosswalk file will be a fixed length text format structured as follows:

1st line

List of variable names as follows:
SSN , BENE_ID



2nd line and beyond

Data for above variables for each beneficiary

The second variable in this file, BENE_ID, refers to the BENE_ID assigned to a beneficiary in the new shipment.

RES_ID / State Code Finder and Crosswalk Files

A RES_ID /State Code finder file must be in CSV (ASCII comma delimited) format structured as follows:

1st line

List of variable names as follows: RES_ID, State Code

2nd line and beyond

Data for above variables for each record:

RES_ID: 10 character alpha-numeric code identifying a resident

State Code: 2 character variable identifying state abbreviation code

The RES_ID / State Code crosswalk file will be a fixed length text format:

1st line

List of variable names as follows:

RES_ID, State Code, BENE_ID

2nd line and beyond

Data for above variables for each beneficiary

The second variable in this file, BENE_ID, refers to the BENE_ID assigned to a beneficiary in the new shipment.

UPIN Finder Files

Submitting a UPIN finder file requires only a listing of UPINs. The following list of UPINS is considered surrogate UPINs and are being used temporarily only while a UPIN is being assigned. These UPINs will be dropped from submitted finder files unless it is specifically requested to keep them.



The following is a list of surrogate UPINs:

<u>Surrogate UPIN</u>	<u>Provider Type</u>
INT000	for each intern
RES000	for each resident
PHS000	for Public Health Service physicians, includes Indian Health Services
VAD000	for Department of Veterans Affairs physicians
RET000	for retired physicians
SLF000	for providers to report that the patient is self-referred
OTH000	for all other unspecified entities not included above.
AA0000	Anesthesia Assistant
CNA000	Certified Registered Nurse Anesthetist
CNM000	Certified Nurse Midwife
CNS000	Clinical Nurse Specialist
CP0000	Clinical Psychologist
CSW000	Clinical Social Worker
FOR000	Foreign Doctor (for all non-United States physicians)
MD0000	Medical Doctor (includes DO, CH, DDM, DDS, DPM, OD)
NP0000	Nurse Practitioner
OT0000	Occupational Therapist
PA0000	Physician Assistant
PT0000	Physical Therapist

A UPIN finder file must be in ASCII format structured as follows:

1st line

List of the single variable name as follows: UPIN

2nd line and beyond

Each UPIN of interest

A carriage return must end each line.