

Files for Order

LDS SAFS: Additional Information

The data is produced in a full DISPLAY format, which means that it is "text ready" and ASCII suitable. It is produced on the mainframe and is kept as EBCDIC text data. When the data is either downloaded or read from a tape reader, it can be read with the EBCDIC/ASCII conversion utility with the CRLF option and will be fully functional on an ASCII platform. The files are immediately readable using NOTEPAD or WORD/EXCEL (ASCII text input). All records in ASCII are of a variable-length nature since ASCII reads to where blanks are at the end of a line and drops the blanks and inserts a CRLF instead. This is the norm for the ASCII world.

The records are of fixed length, but since there may be several different records (Base claim, Diag trailer, line item trailer) all in the same file - the file can be described as a variable file with fixed records inside. When received, any user can merely request that the data be converted from EBCDIC to ASCII with a CRLF (Carriage Return/ Line Feed) and their data will be fully usable on a UNIX or other PC type platform.

Centers for Medicare & Medicaid Services (CMS) uses DESY, which is a CMS Data Extraction System, and when fields are read into the DESY View, if they are packed or signed or decimal fields, these are expanded to a TEXT (DISPLAY) format which means that they are no longer packed, they are given hard signs (+/-) and hard decimal points, so an S9(11)V99 COMP-3 (packed) field is converted to a +9(11).99 field. The original field was 7 bytes and the new field is 15 bytes. The old file would never be able to convert to ASCII text data and could only be read in a binary format. The new field converts successfully to the same 15 bytes (length) in ASCII. This is what makes the data ASCII suitable and why the DESY Views were created the way they were, specifically to support the needs of the users working on an UNIX, NT, or Windows environment.

The changed size of the fields is being documented in the Record Layout for the users.

Questions:

1.) Is the method of selecting beneficiaries for inclusion in the 5% SAF's still the same (i.e. it used to be beneficiaries with 05,20,45,70 or 95 in positions 8 and 9 of the HIC number)

++ YES++ Still the same.

2.) What is the difference between the Beneficiary Encrypted Files (BEFs) and the Limited Data Set (LDS) Files?

++The LDS is the same as our previous BEFs with the exception that the HIC is completely blank in the LDS. All of the other fields are the same as before, except the HIC (HICAN - 11 bytes) is now blank - it is not encrypted anymore, but rather it is replaced with blanks.

The DESY-SORT-KEY is the unique bene identifier field in the LDS files. It will be consistent within the file, across files, and across time, but cannot be used for any other purpose and will not be accepted for search for other CMS data. The DESY-SORT-KEY is provided within 50-bytes generated by DESY and added at the end of the base (or fixed portion) record of each claim. This does not affect the statistical quality of the data as the properties of the data and relative proportions are still the same. In other words, an identified cohort of patients (by diagnosis code as an example) is still the same.

Another difference is the privacy clearance process is less for the LDS files, which should decrease the amount of time involved in receiving the data. Please see the LDS Disclaimer.