

DESCRIPTION OF NCHS-CMS LINKAGE

NCHS has linked NCHS population based survey data with Medicare enrollment and claims data collected from the Centers for Medicare and Medicaid Services (CMS). These linked survey files provide the data needed to formulate and answer vital research questions by profiling Medicare service use and assessing health care costs.

Medicare enrollment and claims data are available for those NCHS respondents who agreed to provide personal identification data to NCHS and for whom NCHS was able to validate and match with Medicare administrative records. CMS provided NCHS with Medicare benefit claims data for 1991 through 2000 for all successfully matched NCHS survey participants.

Important notes regarding analytic issues with the NCHS-CMS linked data files can be found in [Medicare Enrollment and Claims Data Analytic Issues](#).

CMS linked data is available for the following NCHS surveys:

- **1994-1998 National Health Interview Survey (NHIS)**
- **NHANES I Epidemiologic Follow-up Study (NHEFS)**
- **Second National Health and Nutrition Examination Survey (NHANES II)**
- **Third National Health and Nutrition Examination Survey (NHANES III)**
- **The Second Longitudinal Study of Aging (LSOA II)**

The following Medicare enrollment and claims files covering service from 1991-2000 are available for analytic use through the NCHS Research Data Center <http://www.cdc.gov/nchs/r&d/rdc.htm>:

Denominator File

The Denominator file provides data on all Medicare beneficiaries enrolled and/or entitled to Medicare benefits in a given year. Monthly information on the enrollment status of linked Medicare beneficiaries including third party payer information and Group Health Plan (GHP) enrollment information is provided. The Denominator file is fixed length and contains one record per person.

Documentation for the [Denominator File](#) is available in PDF format. The [variable names](#) used in this data file come from the suggested SAS alias variable name provided by CMS in the Denominator file documentation.

MedPAR - Hospital Stay File

The MedPAR Hospital Stay file contains inpatient hospitalization final action claim records. All Medicare Part A short and long stay hospitalization claims for each calendar year are included on the MedPAR Hospital Stay file. Each MedPAR Hospital Stay claim record includes up to 10 ICD-9 diagnoses and 6 ICD-9 procedures associated with each hospital stay. Claim record inclusion on the MedPAR Hospital Stay file is based on year of discharge.

Documentation for the [MedPAR Hospital Stay File](#) is available in PDF format. The [variable names](#) used in this data file come from the suggested standard alias variable names provided by CMS in the MedPAR file documentation.

MedPAR - Skilled Nursing Facility File

The MedPAR Skilled Nursing Facility file contains skilled nursing facility final action claim records. Skilled Nursing Facility (SNF) claims for each calendar year are provided on the MedPAR SNF file. Each MedPAR SNF claim record includes up to 10 ICD-9 diagnoses and 6 ICD-9 procedures provided association with a SNF stay. Inclusion in the MedPAR SNF is based on year of admission into the facility.

Documentation for the [MedPAR Skilled Nursing Facility File](#) is available in PDF format. The [variable names](#) used in this data file come from the suggested standard alias variable names minus the beginning text string “MEDPAR_” in the MedPAR file documentation provided by CMS.

Carrier SAF

The Carrier file (formerly the Physician/Supplier Part B file) contains final action claims data submitted by non-institutional providers. The data is largely made up of physician claim records, although the file also includes claims from other non-institutional providers such as physician assistants, clinical social workers, nurse practitioners, independent clinical laboratories, ambulance providers, and stand-alone ambulatory surgical centers. Due to the large number of carrier claim variables, CMS provides the Carrier data in variable length files. There can be multiple carrier claims per person on a file. The Carrier files are provided in the CMS Standard Analytic File format.

Documentation for the [Carrier SAF](#) is available in PDF format. The [variable names](#) used in this data file come from the suggested SAS alias variable name provided by CMS in the Carrier SAF documentation.

Outpatient SAF

The Outpatient file contains Medicare Part B final action claims from institutional outpatient providers for each calendar year. Hospital outpatient departments, rural health clinics, renal dialysis facilities, outpatient rehabilitation facilities, comprehensive outpatient rehabilitation facilities, and community mental health centers are examples of institutional outpatient providers. Same day surgeries performed in a hospital are also included in the Outpatient file. There can be multiple outpatient claims records per person on the Outpatient files. The Outpatient files are provided in the CMS Standard Analytic File format.

Documentation for the [Outpatient SAF](#) is available in PDF format. The [variable names](#) used in this data file come from the suggested SAS alias variable name provided by CMS in the Outpatient SAF documentation.

Home Health Agency SAF

The Home Health Agency file contains final action claims for home health services. Some of the information contained in this file includes the number of visits, type of visit (skilled-nursing care, home health aides, physical therapy, speech therapy, occupational therapy, and medical social services), diagnosis (10 ICD-9 diagnosis codes), dates of visits, reimbursement amount. An HHA claim may cover services provided over a period of time, not a single day. There can be multiple HHA claims records per person on the HHA files. The Home Health Agency files are provided in the CMS Standard Analytic File format.

Documentation for the [Home Health Agency SAF](#) is available in PDF format. The [variable names](#) used in this data file come from the suggested SAS alias variable names provided by CMS in the Home Health Agency SAF documentation.

Hospice SAF

The Hospice file contains final action claims data submitted by Hospice providers. The data contained in this file include the type of hospice care received (e.g., routine home care, inpatient respite care), The Hospice file contains data fields for 10 ICD-9 diagnosis and 6 procedure codes, dates of service, reimbursement amount, and some demographic information (such as date of birth, race, and sex). There can be multiple Hospice claims records per person on the Hospice file. The Hospice files are provided in the CMS Standard Analytic File format.

Documentation for the [Hospice SAF](#) is available in PDF format. The [variable names](#) used in this data file come from the suggested SAS alias variable names provided by CMS in the Hospice SAF documentation.

Durable Medical Equipment (DMERC)

The DMERC contains final action claims data submitted by Durable Medical Equipment (DME) regional carriers. Some of the information contained in the DMERC includes diagnosis (10 ICD-9 diagnosis codes), service type codes, dates of service, and reimbursement amount. There can be multiple DME claim records per person on the DMERC file.

Documentation for [Durable Medical Equipment](#) is available in PDF format. The [variable names](#) used in this data file come from the suggested SAS alias variable names provided by CMS in the Durable Medical Equipment file documentation.

Summary Medicare Enrollment and Claims (SMEC) Files

The Summary Medicare Enrollment and Claims files were created by NCHS to assist researchers who are interested in analyzing Medicare cost and claims data from multiple Medicare service files. SMEC files are available for each of the NCHS-CMS Medicare linked surveys listed above for each year (1991-2000) of Medicare data. The SMEC files contain data on the beneficiary's reason for Medicare entitlement, total number of months of Medicare entitlement, Group Health Plan (GHP) enrollment, and summarized Medicare service charges and reimbursement amounts. These summarized (or summary) variables are modeled after the [Medicare Current Beneficiary Survey \(MCBS\)](#) cost and use files.

Each SMEC file has a fixed length record format with one record for each survey respondent successfully linked to Medicare administrative record files. Documentation for the [SMEC Files](#) is available in PDF format.

End Stage Renal Disease Data Files

NCHS received a separate set of data files from CMS containing information on patients diagnosed with End Stage Renal Disease (ESRD). The ESRD data files can be used by researchers interested in conducting analysis specifically related to patients with renal dysfunction. The ESRD files are available for each of the NCHS-CMS Medicare linked surveys listed above and contain patient dialysis data from 1975-2000. The ESRD data are divided into three files:

- **Combined ESRD Patient Master and Death Notification File**
- **ESRD Entitlement/Registration File**
- **ESRD Patient Event File**

The Combined ESRD Patient Master and Death Notification files contain general demographic information, primary and contributing causes of ESRD diagnosis, date of first dialysis, and information on kidney transplants. The Combined ESRD Patient Master and Death Notification files also contain death information, if a Death Notification form ([Form 2746](#)) was completed.

Each Combined ESRD Patient Master and Death Notification record has a fixed length format with one record for each successfully linked NCHS survey respondent entered into the CMS ESRD administrative database. Documentation for the [Combined ESRD Patient Master and Death Notification File](#) is available in PDF format.

The ESRD Entitlement/Registration files contain information obtained by CMS from completed ESRD Medical Evidence Report Medicare Entitlement and/or Registration forms ([Form 2728](#)). The ESRD Entitlement/Registration files contains general demographic information, health insurance status, primary and contributing causes of ESRD diagnosis, values from ESRD-related test results, dialysis information, and kidney transplant status.

Each ESRD Entitlement/Registration record has a fixed length format and may contain up to 3 records for each successfully linked NCHS survey respondent entered into the CMS ESRD administrative database. Documentation for the [ESRD Entitlement/Registration File](#) is available in PDF format.

The ESRD Patient Event files contain information regarding ESRD related events, such as dialysis or transplants. The ESRD Patient Event files include on event date, type of event, and patient modality.

The ESRD Patient Event file has a fixed length record format and may contain up to 26 records for each successfully linked NCHS survey respondent entered into the CMS ESRD administrative database. Documentation for the [ESRD Patient Event File](#) is available in PDF format.

Data Usage Issues regarding Public ID/SEQN

The data provided on the 1994-1998 NHIS, NHEFS, NHANES II, NHANES III, and LSOA III linked CMS Medicare files can be merged with the NCHS public use survey data files using the unique survey specific Public Identification number (PUBLIC ID/SEQN). Note: At this time the linked Medicare data files are only available for research use through the NCHS restricted access data center (RDC). Approved RDC researchers may choose to provide their own analytic files created from public use survey files to the RDC. Therefore, it is important for researchers to include survey specific Public Identification number on any analytic files sent to the RDC. The RDC will merge data (using PUBLIC ID or SEQN) from the linked CMS Medicare files to the analyst's file. The merged file will be held at the RDC and made available for analysis. Information on how to identify and/or construct the NCHS survey specific PUBLIC ID or SEQN is provided below.

I. National Health Interview Survey (NHIS)

On the NHIS surveys, researchers need to construct the NHIS public id from the following variables. The number and public-use location varies by NHIS survey year.

NHIS 1994

<u>Item</u>	<u>Public-use Location</u>	<u>Length</u>	<u>Description</u>
Year (2 digit)	3-4	2	Year of interview
Quarter	5	1	Calendar quarter of interview
PSU	6-8	3	Random recode of PSU #
Week	9-10	2	Week of interview within quarter
Segment	11-12	2	Segment number
Household	13-14	2	Household number within quarter
Person number	15-16	2	Person number within household

Note: Concatenate all variables to get the unique person identifier. All variables are zero filled.

NHIS 1995, 1996

<u>Item</u>	<u>Public-use Location</u>	<u>Length</u>	<u>Description</u>
Year (2 digit)	3-4	2	Year of interview
Household ID	5-14	10	Household ID number
Person number	15-16	2	Person number within Household

Note: Concatenate all variables to get the unique person identifier. All variables are zero filled.

NHIS 1997-1998

<u>Item</u>	<u>Location</u>	<u>Length</u>	<u>Description</u>
Year (4 digit)	3-6	4	Year of interview
Household Serial #	7-12	6	Household serial number
Person number	15-16	2	Person number within Household

Note: Concatenate all variables to get the unique person identifier. All variables are zero filled.

II. NHANES I Epidemiologic Follow-up Study NHEFS

<u>Item</u>	<u>Length</u>	<u>Description</u>
SEQN	5	Participant identification number

All of the NHEFS public-use data files are linked with the common survey participant identification number (SEQN). Merging information from multiple NHEFS Files to the NHEFS-CMS linked files using this variable ensures that the appropriate information for each survey participant is linked correctly.

III. Second National Health and Nutrition Examination Survey (NHANES II)

<u>Item</u>	<u>Length</u>	<u>Description</u>
SEQN	5	Participant identification number

All of the NHANES II public-use data files are linked with the common survey participant identification number (SEQN). Merging information from multiple NHANES II Files to the NHANES II-CMS linked files using this variable ensures that the appropriate information for each survey participant is linked correctly.

IV. Third National Health and Nutrition Examination Survey (NHANES III)

<u>Item</u>	<u>Length</u>	<u>Description</u>
SEQN	5	Participant identification number

All of the NHANES III public-use data files are linked with the common survey participant identification number (SEQN). Merging information from multiple NHANES III Files to the NHANES III-CMS linked files using this variable ensures that the appropriate information for each survey participant is linked correctly.

V. The Second Longitudinal Study of Aging (LSOA II)

On the LSOA II survey, researchers need to construct the LSOA II public id from the following variables.

LSOA II

<u>Item</u>	<u>Public-use Location</u>	<u>Length</u>	<u>Description</u>
Year	3-4	2	Year of interview
Quarter	5	1	Calendar quarter of interview
PSU	6-8	3	Random recode of PSU #
Week	9-10	2	Week of interview within quarter
Segment	11-12	2	Segment number
Household	13-14	2	Household number within quarter
Person number	15-16	2	Person number within household

Note: Concatenate all variables to get the unique person identifier. All variables are zero filled.