

NCHS-CMS Medicare Feasibility Files (Updated 11/2016)

NCHS has completed a data linkage project designed to link administrative data collected from the Centers for Medicare & Medicaid Services (CMS) with data collected in several NCHS population surveys.

Medicare enrollment and claims data covering service from **1999-2013** are available for the following NCHS surveys:

- **1994-2013 National Health Interview Survey (NHIS)**
- **1999-2012 National Health and Nutrition Examination Survey (NHANES)**
- **NHANES I Epidemiologic Follow-up Study (NHEFS)**
- **Third National Health and Nutrition Examination Survey (NHANES III)**
- **The Second Longitudinal Study of Aging (LSOA II)**
- **2004 National Nursing Home Survey (NNHS)**
- **2007 National Home and Hospice Care Survey (NHHCS)**

Survey data collected from eligible survey participants for each of these NCHS surveys were linked to CMS Medicare administrative data. The linked CMS Medicare files include: the Master Beneficiary Summary file; the Medicare Provider Analysis and Review (MedPAR) file; the Carrier file; the Home Health Agency (HHA) file; the Durable Medical Equipment (DME) file; the Hospice file; the Outpatient file; and the Part D Prescription Drug Event (PDE) file. More information about the types of data on each of these files is available at the NCHS-CMS Medicare Linked Data website:

<http://www.cdc.gov/nchs/data-linkage/medicare.htm>

Linkage of NCHS survey participants to administrative records is performed for participants who are “linkage-eligible”. Linkage-eligibility for NCHS survey participants is determined by whether participants provided the necessary information: Social Security number (SSN) or Medicare Health Insurance Claim (HIC) number and date of birth or agreed to linkage without SSN or HIC number¹. Linkage-eligibility refers to the potential ability to link data for an NCHS survey participant to CMS Medicare administrative data. It is distinct from program-eligibility, which defines whether an individual meets Federal Medicare eligibility criteria.

NCHS survey participants who were under age 18 at the time of the survey are considered linkage-eligible, if the listed linkage eligibility criteria are met and consent is provided by their parent or guardian. However, in accordance with NCHS Research Ethics Review Board (ERB)² guidelines, NCHS can only provide linked administrative data generated for program participation, claims, and other events that occurred prior to the participant’s 18th birthday.

¹ For 2007-2013 NHIS, a follow up question was asked that allowed NCHS to attempt linkage for participants who indicated that they did not know or refused to provide their SSN or HIC number.

² The NCHS ERB, also known as an Institutional Review Board or IRB, is an administrative body of scientists and non-scientists that is established to protect the rights and welfare of human research subjects.

To maximize public use of these data sets, NCHS has created and released for public use NCHS-CMS Medicare Feasibility data files. These files provide information about survey participant eligibility for CMS linkage (linkage-eligibility), final match status (whether the participants linked to CMS Medicare administrative data), and which specific linked CMS data files are available for each successfully linked survey participant. These files can be used to determine the maximum available sample size so that the feasibility of conducting analyses utilizing the linked CMS Medicare data can be assessed. This should be especially useful to researchers considering whether to initiate a RDC proposal to analyze the restricted use linked CMS Medicare files.

Each feasibility file is NCHS survey and survey year specific. A NCHS-CMS Medicare Feasibility file has been created for all surveys listed above. Each file contains a match status variable: `CMS_MEDICARE_MATCH` -- to indicate whether the NCHS survey participant was eligible for CMS Medicare linkage and if eligible, whether successfully matched to the CMS master enrollment file. The feasibility file also includes a variable to indicate if a survey participant has a linked data record on each of the CMS Medicare administrative record files for each year of Medicare benefit coverage. It is important to note that the variables proposed for this data file do not contain any information about benefits or payments; they are simply designed to give researchers a way to gauge whether their analytic samples will have sufficient numbers of sample persons with linked CMS Medicare data to warrant an application to the RDC.

Due to confidentiality requirements, the NCHS-CMS Medicare linked files are currently restricted for research use and the data are only accessible through the NCHS Research Data Center (RDC), a secure data access portal. Instructions for submitting an RDC proposal to access the data can be found on the [NCHS RDC website](#).

Researchers also may want to consider the percent of participants in a managed care program when determining the feasibility and sample sizes of their proposed research projects. A summary table of the percent of NCHS survey participants who were enrolled in a Medicare Part C plan³ by year and survey can be found at http://www.cdc.gov/nchs/data/datalinkage/cms_medicare_linked_data_medicare_advantage_enrollment_tables.pdf.

³ Medicare Part C plans are also referred to as Medicare Advantage (MA) and include Health Maintenance Organizations (HMOs), Preferred Provider Organizations (PPOs), Private Fee-for-Service (PFFS) Plans, Special Needs Plans, and Medicare Medical Savings Account Plans.

NCHS-CMS Medicare Feasibility Study File DATA CODEBOOK

PUBLICID

NCHS Survey Identifier - Participant Identification Number

Type: Character

Width: 14

File Position: 1-14

Usage Notes:

Public use identifier assigned by NCHS used to link NHIS data and administrative records.

Researchers linking to LSOA II and 1994-2005 NHIS public-use data should use PUBLICID.

See [Appendix A](#) for NCHS survey specific descriptions.

SEQN

NCHS Survey Identifier - Sample Sequence Number

Type: Numeric

Width: 5

File Position: 1-5

Usage Notes:

Public use identifier assigned by NCHS used to link NHANES data and administrative records.

Researchers linking to NHEFS, NHANES III, and 1999-2004 NHANES should use SEQN.

See [Appendix A](#) for NCHS survey specific descriptions.

RESNUM

NCHS Survey Identifier - Resident Record (Case) Number

Type: Numeric

Width: 6

File Position: 1-6

Usage Notes:

Public use identifier assigned by NCHS used to link NCHS survey data and administrative records.

Researchers linking to the 2004 NNHS should use RESNUM.

See [Appendix A](#) for NCHS survey specific descriptions.

NCHS-CMS Medicare Feasibility Files DATA CODEBOOK

ON_MBSF_XXXX[†]

On Master Beneficiary Summary File Indicator (File Year)[†]

Type: Numeric

Width: 1

File Position: 16 - 30

Possible values:

- 0 - Linkage-eligible, linked to one or more years of CMS Medicare administrative data & does not have data on the CMS Medicare Master Beneficiary Summary file for year XXXX[†]
 - Linkage-eligible & not linked to any year of CMS Medicare administrative data
 - Ineligible for CMS Linkage
- 1 Survey participant has data on the CMS Medicare Master Beneficiary Summary file for year XXXX[†]

Usage Notes:

[†]**Represents Medicare coverage years 1999 - 2013**

Survey participants are ineligible for linking to CMS administrative records if they are missing key identification data and/or if they refused to provide their Social Security number (SSN) or Medicare Health Insurance Claim (HIC) number at the time of the survey interview (**for all surveys**) or did not agree to linkage without SSN or HIC number (**for 2007-2013 NHIS**).

In accordance with NCHS Ethics Review Board (ERB) guidelines, NCHS will no longer release linked administrative data for child survey participants if the administrative data was generated for program participation, claims or other events occurring on or after their 18th birthday.

NCHS-CMS Medicare Feasibility Files DATA CODEBOOK

ON_MEDPAR_ XXXX[†]

**On Medicare Provider Analysis and Review (MedPAR) File
Indicator (File Year)[†]**

Type: Numeric

Width: 1

File Position: 31 - 45

Possible values:

- 0** - Linkage-eligible, linked to one or more years of CMS Medicare administrative data & does not have data on the CMS Medicare Provider Analysis and Review (MedPAR) file for year XXXX[†]
- Linkage-eligible & not linked to any year of CMS Medicare administrative data
- Ineligible for CMS Linkage

- 1** Survey participant has data on the CMS Medicare Provider Analysis and Review (MedPAR) file for year XXXX[†]

Usage Notes:

[†]**Represents Medicare coverage years 1999 - 2013**

Survey participants are ineligible for linking to CMS administrative records if they are missing key identification data and/or if they refused to provide their Social Security number (SSN) or Medicare Health Insurance Claim (HIC) number at the time of the survey interview (**for all surveys**) or did not agree to linkage without SSN or HIC number (**for 2007-2013 NHIS**).

In accordance with NCHS Ethics Review Board (ERB) guidelines, NCHS will no longer release linked administrative data for child survey participants if the administrative data was generated for program participation, claims or other events occurring on or after their 18th birthday.

NCHS-CMS Medicare Feasibility Files DATA CODEBOOK

ON_CARRIER_XXXX[†]

On Carrier File Indicator (File Year)[†]

Type: Numeric

Width: 1

File Position: 46 - 60

Possible values:

- 0** - Linkage-eligible, linked to one or more years of CMS Medicare administrative data & does not have data on the CMS Medicare Carrier file for year XXXX[†]
 - Linkage-eligible & not linked to any year of CMS Medicare administrative data
 - Ineligible for CMS Linkage

- 1** Survey participant has data on the CMS Medicare Carrier file for year XXXX[†]

Usage Notes:

[†]**Represents Medicare coverage years 1999 – 2013**

Survey participants are ineligible for linking to CMS administrative records if they are missing key identification data and/or if they refused to provide their Social Security number (SSN) or Medicare Health Insurance Claim (HIC) number at the time of the survey interview (**for all surveys**) or did not agree to linkage without SSN or HIC number (**for 2007-2013 NHIS**).

In accordance with NCHS Ethics Review Board (ERB) guidelines, NCHS will no longer release linked administrative data for child survey participants if the administrative data was generated for program participation, claims or other events occurring on or after their 18th birthday.

NCHS-CMS Medicare Feasibility Files DATA CODEBOOK

ON_DME_XXXX[†]

On Durable Medical Equipment (DME) File Indicator (File Year)[†]

Type: Numeric

Width: 1

File Position: 61 - 75

Possible values:

- 0 - Linkage-eligible, linked to one or more years of CMS Medicare administrative data & does not have data on the CMS Medicare Durable Medical Equipment (DME) file for year XXXX[†]
 - Linkage-eligible & not linked to any year of CMS Medicare administrative data
 - Ineligible for CMS Linkage
- 1 Survey participant has data on the CMS Medicare Durable Medical Equipment (DME) file for year XXXX[†]

Usage Notes:

[†]**Represents Medicare coverage years 1999 – 2013**

Survey participants are ineligible for linking to CMS administrative records if they are missing key identification data and/or if they refused to provide their Social Security number (SSN) or Medicare Health Insurance Claim (HIC) number at the time of the survey interview (**for all surveys**) or did not agree to linkage without SSN or HIC number (**for 2007-2013 NHIS**).

In accordance with NCHS Ethics Review Board (ERB) guidelines, NCHS will no longer release linked administrative data for child survey participants if the administrative data was generated for program participation, claims or other events occurring on or after their 18th birthday.

NCHS-CMS Medicare Feasibility Files DATA CODEBOOK

ON_HHA_XXXX[†]

On Home Health Agency (HHA) File Indicator (File Year)[†]

Type: Numeric

Width: 1

File Position: 76 - 90

Possible values:

- 0** - Linkage-eligible, linked to one or more years of CMS Medicare administrative data & does not have data on the CMS Medicare Home Health Agency (HHA) file for year XXXX[†]
 - Linkage-eligible & not linked to any year of CMS Medicare administrative data
 - Ineligible for CMS Linkage
- 1** Survey participant has data on the CMS Medicare Home Health Agency (HHA) file for year XXXX[†]

Usage Notes:

[†]**Represents Medicare coverage years 1999 – 2013**

Survey participants are ineligible for linking to CMS administrative records if they are missing key identification data and/or if they refused to provide their Social Security number (SSN) or Medicare Health Insurance Claim (HIC) number at the time of the survey interview (**for all surveys**) or did not agree to linkage without SSN or HIC number (**for 2007-2013 NHIS**).

In accordance with NCHS Ethics Review Board (ERB) guidelines, NCHS will no longer release linked administrative data for child survey participants if the administrative data was generated for program participation, claims or other events occurring on or after their 18th birthday.

NCHS-CMS Medicare Feasibility Files DATA CODEBOOK

ON_HOSPICE_XXXX[†]

On Hospice File Indicator (File Year)[†]

Type: Numeric

Width: 1

File Position: 91 - 105

Possible values:

- 0 - Linkage-eligible, linked to one or more years of CMS Medicare administrative data & does not have data on the CMS Medicare Hospice file for year XXXX[†]
 - Linkage-eligible & not linked to any year of CMS Medicare administrative data
 - Ineligible for CMS Linkage
- 1 Survey participant has data on the CMS Medicare Hospice file for year XXXX[†]

Usage Notes:

[†]Represents Medicare coverage years 1999 – 2013

Survey participants are ineligible for linking to CMS administrative records if they are missing key identification data and/or if they refused to provide their Social Security number (SSN) or Medicare Health Insurance Claim (HIC) number at the time of the survey interview (**for all surveys**) or did not agree to linkage without SSN or HIC number (**for 2007-2013 NHIS**).

In accordance with NCHS Ethics Review Board (ERB) guidelines, NCHS will no longer release linked administrative data for child survey participants if the administrative data was generated for program participation, claims or other events occurring on or after their 18th birthday.

NCHS-CMS Medicare Feasibility Files DATA CODEBOOK

ON_OUTPAT_XXXX[†]

On Outpatient File Indicator (File Year)[†]

Type: Numeric

Width: 1

File Position: 106 - 120

Possible values:

- 0** - Linkage-eligible, linked to one or more years of CMS Medicare administrative data & does not have data on the CMS Medicare Outpatient file for year XXXX[†]
- Linkage-eligible & not linked to any year of CMS Medicare administrative data
- Ineligible for CMS Linkage

- 1** - Survey participant has data on the CMS Medicare Outpatient file for year XXXX[†]

Usage Notes:

[†]**Represents Medicare coverage years 1999 – 2013**

Survey participants are ineligible for linking to CMS administrative records if they are missing key identification data and/or if they refused to provide their Social Security number (SSN) or Medicare Health Insurance Claim (HIC) number at the time of the survey interview (**for all surveys**) or did not agree to linkage without SSN or HIC number (**for 2007-2013 NHIS**).

In accordance with NCHS Ethics Review Board (ERB) guidelines, NCHS will no longer release linked administrative data for child survey participants if the administrative data was generated for program participation, claims or other events occurring on or after their 18th birthday.

NCHS-CMS Medicare Feasibility Files DATA CODEBOOK

ON_PARTD_PDE_XXXX‡

On Part D Prescription Drug Event (PDE) File Indicator (File Year)‡

Type: Numeric

Width: 1

File Position: 121 - 128

Possible values:

- 0 - Linkage-eligible, linked to one or more years of CMS Medicare administrative data & does not have data on the Part D Prescription Drug Event (PDE) file for year XXXX‡
 - Linkage-eligible & not linked to any year of CMS Medicare administrative data
 - Ineligible for CMS Linkage
- 1 Survey participant has data on the Part D Prescription Drug Event (PDE) file for year XXXX‡

Usage Notes:

‡Represents Medicare coverage years 2006-2013

Survey participants are ineligible for linking to CMS administrative records if they are missing key identification data and/or if they refused to provide their Social Security number (SSN) or Medicare Health Insurance Claim (HIC) number at the time of the survey interview (**for all surveys**) or did not agree to linkage without SSN or HIC number (**for 2007-2013 NHIS**).

In accordance with NCHS Ethics Review Board (ERB) guidelines, NCHS will no longer release linked administrative data for child survey participants if the administrative data was generated for program participation, claims or other events occurring on or after their 18th birthday.

Appendix A: Data Usage Issues regarding PUBLICID/SEQN/RESNUM

The data provided on the 1994-2013 NHIS, NHEFS, 1999-2012 NHANES, NHANES III, LSOA II, 2004 NNHS, and the 2007 NHHCS linked CMS Medicare files can be merged with the NCHS public use survey data files using the unique survey specific Public Identification number (PUBLICID/SEQN/RESNUM/PATNUM).

Note: At this time the linked CMS 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 PUBLICID, SEQN, RESNUM or PATNUM) 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 PUBLICID, SEQN, RESNUM or PATNUM is provided below.

I. National Health Interview Survey (NHIS)

NHIS 1994

<u>Variable</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
PSUNUMR	6-8	3	Random recode of PSU
WEEKCEN	9-10	2	Week of interview within quarter
SEGNUM	11-12	2	Segment number
HHNUM	13-14	2	Household number within quarter
PNUM	15-16	2	Person number within household

Note: Concatenate all variables to get the unique person identifier.

SAS example:

```
length publicid $14;  
PUBLICID = trim(left(YEAR | QUARTER | PSUNUMR | WEEKCEN | SEGNUM | HHNUM | PNUM));
```

Stata example: (note this will convert the variables to string variables)

```
egen PUBLICID = concat(YEAR QUARTER PSUNUMR WEEKCEN SEGNUM HHNUM PNUM)
```

NHIS 1995, 1996

<u>Variable</u>	<u>Public-use Location</u>	<u>Length</u>	<u>Description</u>
YEAR	3-4	2	Year of interview
HHID	5-14	10	Household ID number
PNUM	15-16	2	Person number within household

Note: Concatenate all variables to get the unique person identifier.

SAS example:

```
length publicid $14;  
PUBLICID = trim(left(YEAR || HHID || PNUM));
```

Stata example: (note this will convert the variables to string variables)

```
egen PUBLICID = concat(YEAR HHID PNUM)
```

NHIS 1997-2003

<u>Variable</u>	<u>Public-use Location</u>	<u>Length</u>	<u>Description</u>
SRVY_YR	3-6	4	Year of interview
HHX	7-12	6	Household number
FMX	13-14	2	Family number
PX	15-16	2	Person number within household

Note: Concatenate all variables to get the unique person identifier.

SAS example:

```
length publicid $14;  
PUBLICID = trim(left(SRVY_YR || HHX || FMX || PX));
```

Stata example: (note this will convert the variables to string variables)

```
egen PUBLICID = concat(SRVY_YR HHX FMX PX)
```

*The person identifier was called PX in the 1997-2003 NHIS and FPX in the 2004 (and later) NHIS; users may find it necessary to create an FPX variable in the 2003 and earlier datasets (or PX in later datasets).

NHIS 2004

<u>Variable</u>	<u>Location</u>	<u>Length</u>	<u>Description</u>
SRVY_YR	3-6	4	Year of interview
HHX	7-12	6	Household number
FMX	13-14	2	Family number
FPX	15-16	2	Person number within household

Note: Concatenate all variables to get the unique person identifier.

SAS example:

```
length publicid $14;  
PUBLICID = trim(left(SRVY_YR || HHX || FMX || FPX));
```

Stata example: (note this will convert the variables to string variables)

```
egen PUBLICID = concat(SRVY_YR HHX FMX FPX)
```

NHIS 2005-2013

<u>Variable</u>	<u>Location</u>	<u>Length</u>	<u>Description</u>
SRVY_YR	3-6	4	Year of interview
HHX	7-12	6	Household number
FMX	16-17	2	Family number
FPX	18-19	2	Person number within household

Note: Concatenate all variables to get the unique person identifier.

SAS example:

```
length publicid $14;  
PUBLICID = trim(left(SRVY_YR || HHX || FMX || FPX));
```

Stata example: (note this will convert the variables to string variables)

```
egen PUBLICID = concat(SRVY_YR HHX FMX FPX)
```

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 Medicare linked files using this variable ensures that the appropriate information for each survey participant is linked correctly.

III. National Health and Nutrition Examination Survey (NHANES)

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

All of the NHANES public-use data files are linked with the common survey participant identification number (SEQN). Merging information from multiple NHANES Files to the NHANES-CMS Medicare 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 Medicare 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>Variable</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
PSUNUMR	6-8	3	Random recode of PSU #
WEEKCEN	9-10	2	Week of interview within quarter
SEGNUM	11-12	2	Segment number
HHNUM	13-14	2	Household number within quarter
PNUM	15-16	2	Person number within household

Note: Concatenate all variables to get the unique person identifier.

SAS example:

```
length publicid $14;  
PUBLICID = trim(left(YEAR | QUARTER | PSUNUMR | WEEKCEN | SEGNUM | HHNUM | PNUM) );
```

Stata example: (note this will convert the variables to string variables)

```
egen PUBLICID = concat(YEAR QUARTER PSUNUMR WEEKCEN SEGNUM HHNUM PNUM)
```


VI. 2004 National Nursing Home Survey (NNHS)

<u>Item</u>	<u>Length</u>	<u>Description</u>
RESNUM	6	Resident Record (Case) Number

All of the 2004 NNHS public-use data files are linked with the common resident record (case) number (RESNUM). Merging information from the 2004 NNHS Files to the 2004 NNHS-CMS Medicare linked files using this variable ensures that the appropriate information for each survey participant is linked correctly.

VII. 2007 National Home and Hospice Care Survey (NHHCS)

<u>Item</u>	<u>Length</u>	<u>Description</u>
PATNUM	6	Patient/Discharge Record (Case) Number

All of the 2007 NHHCS public-use data files are linked with the common patient/discharge record (case) number (PATNUM). Merging information from the 2007 NHHCS Files to the 2007 NHHCS-CMS Medicare linked files using this variable ensures that the appropriate information for each survey participant is linked correctly.