# Linkage File for 2023 MEPS and 2021-2022 NHIS Public Use Files

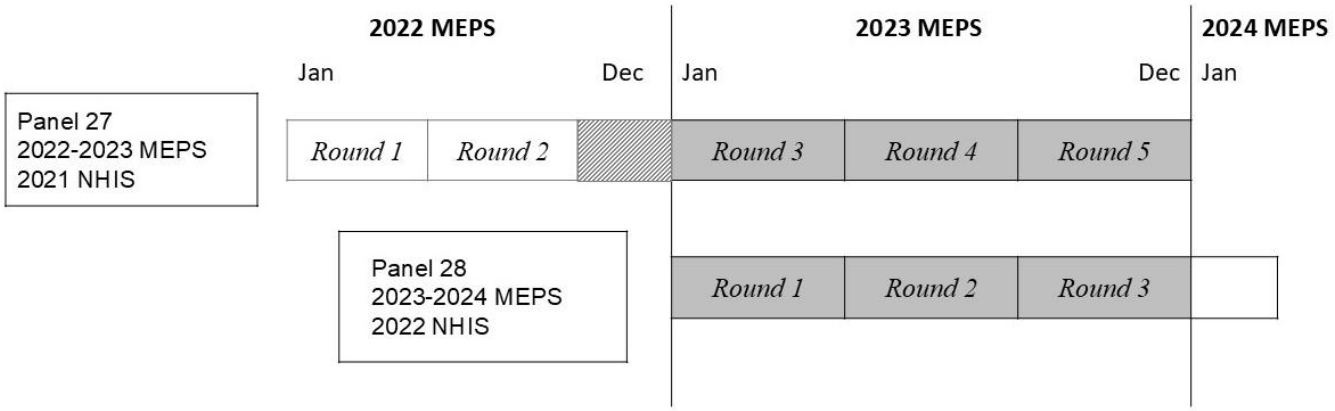
## 1.0 Overview

The Medical Expenditure Panel Survey (MEPS) Household Component (HC) uses the National Health Interview Survey (NHIS) as its sampling frame. Each year a new MEPS-HC panel is established by drawing the sample from the previous year's NHIS responding sample. The MEPS-HC design is based on an overlapping panel design, and two consecutive panels are overlapped in 2023 MEPS full year (FY) file. Details of the 2023 panel design are provided in the documentation for the 2023 Full Year Population Characteristic File.[1]

Due to the 2019 NHIS redesign, from each household only one adult and one child (if any children live in the household) were included in the NHIS sample.[2] As a result, potentially only one adult and one child (if any), in each MEPS household could be linked to the previous data year's NHIS public use file. However, prior to the 2019 NHIS redesign, all persons in an NHIS household were included in the NHIS PUF that made it possible to potentially link nearly all persons in a MEPS panel.[3]

As illustrated in Figure 1 below, the 2023 MEPS full-year public use files (PUFs) cover the calendar year and contain data from Rounds 3, 4, and 5 of MEPS Panel 27 (which uses the 2021 NHIS as its sampling frame), and Rounds 1, 2, and 3 of MEPS Panel 28 (which uses the 2022 NHIS as its sampling frame). Therefore, MEPS Panels 27 and 28 can be linked to 2021 and 2022 NHIS PUFs respectively.

**Figure 1. Mapping of MEPS Year, Panels, and Rounds to NHIS Years**




PUFs containing NHIS data for a given calendar year are available from the National Center for Health Statistics (NCHS).

Users who need to augment the MEPS data with information from NHIS can do so with the linkage file described in the following sections.

## 2.0 Linkage File Description

The MEPS and NHIS linkage file, NHMEP23X.DAT, allows the data user to merge any of the person-

1 https://meps.ahrq.gov/data_stats/download_data/pufs/h247/h247doc.pdf
2 For details of the 2019 redesign, see https://www.cdc.gov/nchs/nhis/2019_quest_redesign.htm.
3 A small number of persons may not be linkable between NHIS and MEPS because they may have left or joined the household between when NHIS was fielded and when MEPS was fielded.

level 2023 MEPS full-year public use data files with the 2021 and 2022 NHIS person-level PUFs (Sample Adult and Sample Child).

The NHIS person identifiers changed in 2019. Prior to 2019, each family (FMX) has been considered a separate case, and unique person identifiers have been Household Serial Number (HHX), Family Sequence Number (FMX), and Person Sequence Number (FPX). Beginning in 2019, only a sample adult and, where available, a sample child were included from each household. Therefore, the identifiers in 2019 and later are HHX and record type (RECTYPE) that specifies sample adult, sample child, or not sampled for NHIS.

The linkage file contains 18,919 person-level records and six variables. In the linkage file, a record exists for each of the MEPS 2023 full-year persons. Each record contains the MEPS sample person ID (DUPERSID) and the corresponding NHIS sample person IDs (HHX and RECTYPE). The linkage file can be linked to any of the person-level MEPS 2023 full-year public use data files using the variable DUPERSID. The linkage file can be linked to the NHIS 2021 or 2022 sample adult and sample child data files by HHX, RECTYPE, and SRVY_YR.

When a MEPS sample person does not link to NHIS, HHX is set to 9999999, SRVY_YR is set to 9999, RECTYPE is set to 99, and LINKFLAG is set to 0.

## 3.0 Linkage File Record Counts

Of the 8,434 MEPS Panel 27 persons, 4,062 persons link to the 2021 NHIS data; 5,117 of the 10,485 Panel 28 persons link to the 2022 NHIS data. A total of 9,740 persons in the two panels do not link to either 2021 or 2022 NHIS data. These unlinked cases include newborns; newly in-scope persons; a small number of cases where the NHIS identified a household as responding, but when fielded in MEPS it was determined to actually be a nonresponding household; and household members who are neither the sample adult nor sample child. Table 1 below summarizes the linkages.

**Table 1. Linkage File Record Counts**

| 2023 MEPS Full-Year Data | Total 2023 MEPS Persons | Linked to 2021 or 2022 NHIS PUF (total observations in NHIS PUF) | Not Linked to NHIS |
|---|---|---|---|
| **Panel 27 persons (2021 NHIS)** | 8,434 | 4,062 (37,743) | 4,372 |
| **Panel 28 persons (2022 NHIS)** | 10,485 | 5,117 (35,115) | 5,368 |
| **Total** | 18,919 | 9,179 (72,858) | 9,740 |

## 4.0 Linkage File Record Layout

Table 2 is the record layout for the person-level MEPS-NHIS linkage file (NHMEP23X.DAT).

**Table 2. Layout for the Person-Level MEPS-NHIS Linkage File**

| Variable | Columns | Type | Label and value range* |
|---|---|---|---|
| | | | |

| Variable | Columns | Type | Label and value range* |
|---|---|---|---|
| **DUPERSID** | 1–10 | Character | MEPS encrypted person ID (range = 2790002101 – 2819793102) |
| **HHX** | 11–17 | Character | NHIS household serial number (range = H000014 – H067777) |
| **LINKFLAG** | 18–18 | Numeric | Linkage status between MEPS and NHIS (1 or 0) |
| **PANEL** | 19–20 | Numeric | MEPS panel number (27 or 28) |
| **SRVY_YR** | 21–24 | Numeric | NHIS survey year (2021 or 2022) |
| **RECTYPE** | 25-26 | Numeric | Record type (10 or 20) |

*Values may be missing based on NHIS survey year or linkage status.

Below is the input statement to convert the linkage file (NHMEP23X.DAT) to a SAS dataset.

```
DATA XX.NHMEP23X;
    INFILE    "C:\TEMP\MEPS\NHMEP23X.DAT";
     INPUT  DUPERSID $1-10  HHX $11-17  LINKFLAG 18  PANEL  19-20 SRVY_YR 21-24 RECTYPE 25-26;
RUN;
```

## 5.0 Linking Instructions for SAS Users

The following is one way of adding NHIS person-level variables to the MEPS person-level file. Input files are MEPS HC-247 (2023 Full-Year Population Characteristic File), the 2021 NHIS sample adult and sample child data files, the 2022 NHIS sample adult and sample child data files, and the linkage file NHMEP23X.DAT.

(1) Create seven SAS datasets as follows:

- Convert MEPS HC-247 (ASCII, SAS transport file, or SAS V9 file) to a SAS dataset named FY2023 (n =18,919).
- Convert the linkage file NHMEP23X.DAT to a SAS dataset named NHMEP23X (n =18,919).
- Convert the 2021 NHIS Sample Adult file to a SAS dataset named NHIS2021A (n = 29,482). Make sure the SAS dataset includes HHX, RECTYPE, SRVY_YR, and other variables that are to be added to the MEPS full-year dataset.
- Convert the 2021 NHIS Sample Child file to a SAS dataset named NHIS2021C (n = 8,261). Make sure the SAS dataset includes HHX, RECTYPE, SRVY_YR, and other variables that are to be added to the MEPS full-year dataset.
- Convert the 2022 NHIS Sample Adult file to a SAS dataset named NHIS2022A (n = 27,651). Make sure the SAS dataset includes HHX, RECTYPE, SRVY_YR, and other variables that are to be added to the MEPS full-year dataset.
- Convert the 2022 NHIS Sample Child file to a SAS dataset named NHIS2022C (n = 7,464). Make sure the SAS dataset includes HHX, RECTYPE, SRVY_YR, and other variables that are to be added to the MEPS full-year dataset.

(2) Sort FY2023 by DUPERSID. Concatenate NHIS2021A, NHIS2021C, NHIS2022A, and NHIS2022C into one dataset named NHISALL (n = 72,858). Sort NHISALL by HHX, RECTYPE and SRVY_YR.

(3) Merge FY2023 (n =18,919) with NHMEP23X (n = 18,919) by DUPERSID. Name the output dataset MEPS (n = 18,919). Then sort MEPS by HHX, RECTYPE and SRVY_YR.

(4) Merge MEPS (n = 18,919) with NHISALL (n = 72,858) by HHX, RECTYPE and SRVY_YR. Keep records only in MEPS (n = 18,919). Name the output dataset MEPS23NH (n = 18,919).

**Sample SAS Code for Adding NHIS Variables to the MEPS Dataset**.

```
LIBNAME MEPS "C:\TEMP\MEPS";    /*MEPS 2023 Full-Year PUF, MEPS-NHIS Link, output file*/
LIBNAME NHIS "C:\TEMP\NHIS";    /*NHIS 2021 and 2022 Sampled Adult and Child Files*/

PROC FORMAT;
 VALUE AGE
 .='.'
 0-HIGH='>=0';
RUN;

PROC SORT DATA=MEPS.FY2023;
 BY DUPERSID;
RUN;

DATA NHISALL;
  SET    NHIS.NHIS2021A (KEEP=HHX RECTYPE SRVY_YR AGEP_A RENAME=(AGEP_A=AGE_P) /*other NHIS variables*/)
         NHIS.NHIS2021C (KEEP=HHX RECTYPE SRVY_YR AGEP_C RENAME=(AGEP_C=AGE_P) /*other NHIS variables*/)
         NHIS.NHIS2022A (KEEP=HHX RECTYPE SRVY_YR AGEP_A RENAME=(AGEP_A=AGE_P) /*other NHIS variables*/)
         NHIS.NHIS2022C (KEEP=HHX RECTYPE SRVY_YR AGEP_C RENAME=(AGEP_C=AGE_P) /*other NHIS variables*/);
       FORMAT _ALL_;

RUN;

PROC SORT DATA=NHISALL;
 BY HHX RECTYPE SRVY_YR;
RUN;

DATA MEPS;
  MERGE MEPS.FY2023 MEPS.NHMEP23X (KEEP=DUPERSID HHX RECTYPE SRVY_YR LINKFLAG);
  BY DUPERSID;
RUN;

PROC SORT DATA=MEPS;
 BY HHX RECTYPE SRVY_YR;
RUN;

DATA MEPS.MEPS23NH;
  MERGE MEPS (IN=A) NHISALL;
  BY HHX RECTYPE SRVY_YR;
  IF A;
RUN;

TITLE1 "MEPS 2023 FY data with NHIS variables";
PROC FREQ DATA=MEPS.MEPS23NH;
 TABLES LINKFLAG*SRVY_YR*AGE_P/LIST MISSING;
 FORMAT AGE_P AGE.;
RUN;
```

### Sample Stata Code for Adding NHIS Variables to the MEPS Dataset

```
cd "c:temp"

log using stata23.log, replace

use "meps\h247", clear
rename *, lower
sort dupersid

tempfile fy2023
save `fy2023', replace

use "nhis\nhis2021a", clear
append using "nhis\nhis2021c"
append using "nhis\nhis2022a"
append using "nhis\nhis2022c"

rename *, lower
sort hhx rectype srvy_yr
tempfile nhisall
save `nhisall', replace

infix str dupersid 1-10 str hhx 11-17 linkflag 18 panel 19-20 srvy_yr 21-24 rectype 25-26 using "meps\nhmep23x.dat", clear
sort dupersid
tempfile link
save `link', replace

use `fy2023'
merge 1:1 dupersid using `link'
drop _merge
sort hhx rectype srvy_yr
tempfile meps
save `meps', replace

merge m:1 hhx rectype srvy_yr using `nhisall'
keep if _merge != 2 /*drop cases where a record was found in the NHIS PUFs but not in MEPS*/
keep dupersid hhx rectype srvy_yr linkflag /*edit this line to add any other desired nhis variables*/
save "meps\meps23nh", replace

describe
tab srvy_yr linkflag, missing

log close
```

## 6.0 Further Information

For any questions regarding the linkage file, please contact Anita Soni at Anita.Soni@ahrq.hhs.gov. MEPS public use data files can be downloaded free of charge from the MEPS website at https://www.meps.ahrq.gov. NHIS public use data files can be obtained by contacting NCHS by telephone (301–458–4636) or through their website, https://www.cdc.gov/nchs.