

Linkage File for 2022 MEPS and 2018, 2020, and 2021 NHIS Public Use Files

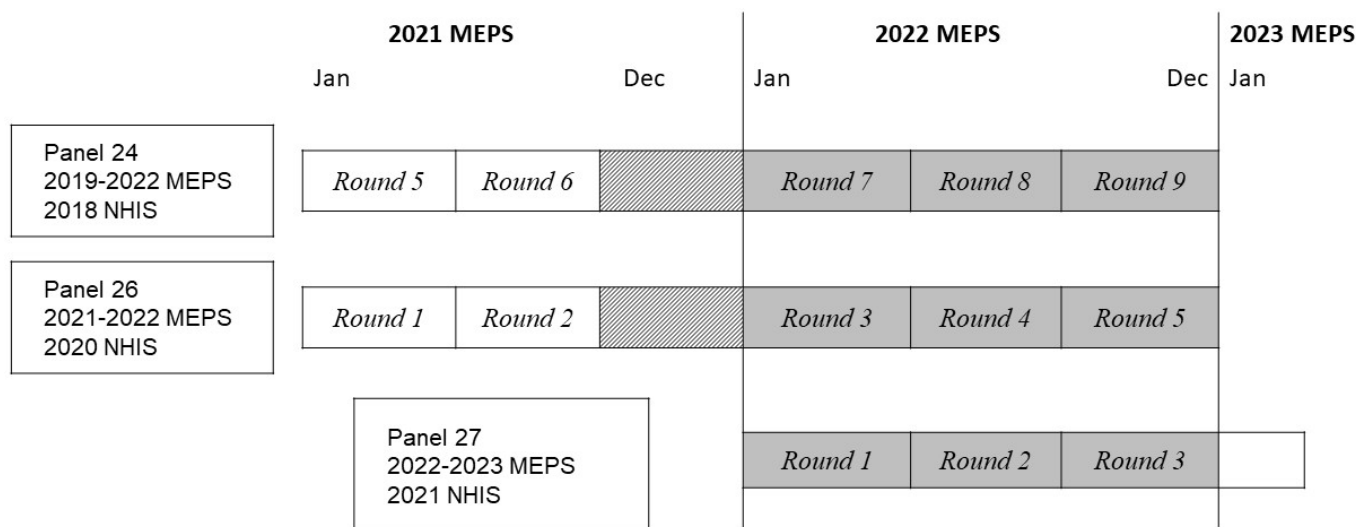
1.0 Overview

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. Usually, two consecutive panels are overlapped in a MEPS full year (FY) file. However, the 2022 FY file included three overlapping panels, Panels 24, 26 and 27, to increase the sample size to compensate for low response rates and data collection difficulties due to the coronavirus 2019 (COVID-19) pandemic. Details of the 2022 panel design are provided in the [documentation for the 2022 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 2022 MEPS full-year public use files (PUFs) cover the calendar year and contain data from Rounds 7, 8 and 9 of MEPS Panel 24 (which uses the 2018 NHIS as its sampling frame), Rounds 3, 4, and 5 of MEPS Panel 26 (which uses the 2020 NHIS as its sampling frame), and Rounds 1, 2, and 3 of MEPS Panel 27 (which uses the 2021 NHIS as its sampling frame). Therefore, MEPS Panels 24, 26, and 27 can be linked to 2018, 2020, and 2021 NHIS PUFs respectively.

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



¹ https://meps.ahrq.gov/data_stats/download_data/pufs/h238/h238doc.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.

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, NHMEP22X.DAT, allows the data user to merge any of the person-level 2022 MEPS full-year public use data files with the 2018, 2020, and 2021 NHIS person-level PUFs (Person, 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 22,431 person-level records and eight variables. In the linkage file, a record exists for each of the MEPS 2022 full-year persons. Each record contains the MEPS sample person ID (DUPERSID) and the corresponding NHIS sample person IDs (HHX, FMX, FPX, and RECTYPE). The linkage file can be linked to any of the person-level MEPS 2022 full-year public use data files using the variable DUPERSID. The linkage file can be linked to the NHIS 2018 person-level data files by HHX, FMX, FPX, and SRVY_YR and to the NHIS 2020 or 2021 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 999999, FMX is set to 99, FPX is set to 99, 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 5,437 MEPS Panel 24 persons, 4,928 persons link to the 2018 NHIS data; 3,435 of the 6,843 Panel 26 persons link to the 2020 NHIS data, and 4,807 of the 10,151 Panel 27 persons link to the 2021 NHIS data. A total of 9,261 persons in the three panels do not link to either 2018, 2020, or 2021 NHIS data. For Panel 24, these unlinked cases include newborns; newly in-scope persons; and 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. As mentioned above, starting with Panel 25, unlinked cases may also be household members who are neither the sample adult nor sample child. Table 1 below summarizes the linkages.

Table 1. Linkage File Record Counts

2022 MEPS Full-Year Data	Total 2022 MEPS Persons	Linked to 2018, 2020 or 2021 NHIS PUF (total observations in NHIS PUF)	Not Linked to NHIS
Panel 24 persons (2018 NHIS)	5,437	4,928 (72,831)	509
Panel 26 persons (2020 NHIS)	6,843	3,435 (37,358)	3,408

2022 MEPS Full-Year Data	Total 2022 MEPS Persons	Linked to 2018, 2020 or 2021 NHIS PUF (total observations in NHIS PUF)	Not Linked to NHIS
Panel 27 persons (2021 NHIS)	10,151	4,807 (37,743)	5,344
Total	22,431	13,170 (147,932)	9,261

4.0 Linkage File Record Layout

Table 2 is the record layout for the person-level MEPS-NHIS linkage file (NHMEP22X.DAT).

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

Variable	Columns	Type	Label and value range*
DUPERSID	1–10	Character	MEPS encrypted person ID (range = 2460002101 – 2799700101)
HHX	11–17	Character	NHIS household serial number (range = 00003817 – H070028)
FMX	18–19	Character	NHIS family number (range = 01 – 06)
FPX	20–21	Character	NHIS person number (range = 01 – 10)
LINKFLAG	22–22	Numeric	Linkage status between MEPS and NHIS (1 or 0)
PANEL	23–24	Numeric	MEPS panel number (24, 26 or 27)
SRVY_YR	25–28	Numeric	NHIS survey year (2018, 2020, or 2021)
RECTYPE	29-30	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 (NHMEP22X.DAT) to a SAS dataset.

```
DATA XX.NHMEP22X;
  INFILE "C:\TEMP\MEPS\NHMEP22X.DAT";
  INPUT DUPERSID $1-10 HHX $11-17 FMX $18-19 FPX $20-21 LINKFLAG 22 PANEL 23-24 SRVY_YR 25-28
RECTYPE 29-30;
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-238 (2022 Full-Year Population Characteristic File), the 2018 NHIS person-level data file, the 2020 NHIS sample adult and sample child data files, the 2021 NHIS sample adult and sample child data files, and the linkage file NHMEP22X.DAT.

- (1) Create seven SAS datasets as follows:
 - Convert MEPS HC-238 (ASCII, SAS transport file, or SAS V9 file) to a SAS dataset named FY2022 (n =22,431).
 - Convert the linkage file NHMEP22X.DAT to a SAS dataset named NHMEP22X (n =22,431).

- Convert the 2018 NHIS Person file to a SAS dataset named NHIS2018 (n = 72,831). Make sure the SAS dataset includes HHX, FMX, FPX, RECTYPE, SRVY_YR, and other variables that are to be added to the MEPS full-year dataset.
 - Convert the 2020 NHIS Sample Adult file to a SAS dataset named NHIS2020A (n = 31,568). 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 2020 NHIS Sample Child file to a SAS dataset named NHIS2020C (n = 5,790). 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 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.
- (2) Sort FY2022 by DUPERSID. Concatenate NHIS2018, NHIS2020A, NHIS2020C, NHIS2021A, and NHIS2021C into one dataset named NHISALL (n = 147,932). Sort NHISALL by HHX, FMX, FPX, RECTYPE and SRVY_YR.
- (3) Merge FY2022 (n = 22,431) with NHMEP22X (n = 22,431) by DUPERSID. Name the output dataset MEPS (n = 22,431). Then sort MEPS by HHX, FMX, FPX, RECTYPE and SRVY_YR.
- (4) Merge MEPS (n = 22,431) with NHISALL (n = 147,932) by HHX, FMX, FPX, RECTYPE and SRVY_YR. Keep records only in MEPS (n = 22,431). Name the output dataset MEPS22NH (n = 22,431).

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

```
LIBNAME MEPS "C:\TEMP\MEPS"; /*MEPS 2022 Full-Year PUF, MEPS-NHIS Link, output file*/
LIBNAME NHIS "C:\TEMP\NHIS"; /*NHIS 2018 Person File and NHIS 2020 and 2021 Sampled Adult and Child Files*/

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

PROC SORT DATA=MEPS.FY2022;
  BY DUPERSID;
RUN;

DATA NHISALL;
  SET NHIS.NHIS2018 (KEEP=HHX FMX FPX RECTYPE SRVY_YR AGE_P /*other NHIS variables*/)
    NHIS.NHIS2020A (KEEP=HHX RECTYPE SRVY_YR AGEP_A RENAME=(AGEP_A=AGE_P) /*other NHIS variables*/)
    NHIS.NHIS2020C (KEEP=HHX RECTYPE SRVY_YR AGEP_C RENAME=(AGEP_C=AGE_P) /*other NHIS variables*/);
  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*/);

RUN;
```

```

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

DATA MEPS;
  MERGE MEPS.FY2022 MEPS.NHMEP22X (KEEP=DUPERSID HHX FMX FPX RECTYPE SRVY_YR LINKFLAG);
  BY DUPERSID;
RUN;

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

DATA MEPS.MEPS22NH;
  MERGE MEPS (IN=A) NHISALL;
  BY HHX FMX FPX RECTYPE SRVY_YR;
  IF A;
RUN;

TITLE1 "MEPS 2022 FY data with NHIS variables";
PROC FREQ DATA=MEPS.MEPS22NH;
  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 stata22.log, replace

use "meps\h238", clear
rename *, lower
sort dupersid

tempfile fy2022
save `fy2022', replace

use "nhis\nhis2018", clear
append using "nhis\nhis2020a"
append using "nhis\nhis2020c"
append using "nhis\nhis2021a"
append using "nhis\nhis2021c"

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

infix str dupersid 1-10 str hhx 11-17 str fmx 18-19 str fpx 20-21 linkflag 22 panel 23-24 srvy_yr 25-28 rectype 29-30 using
"meps\nhmep22x.dat", clear
sort dupersid
tempfile link
save `link', replace

use `fy2022'
merge 1:1 dupersid using `link'
drop _merge
sort hhx fmx fpx rectype srvy_yr
tempfile meps

```

```
save `meps`, replace
```

```
merge m:1 hhx fmx fpx 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 fmx fpx rectype srvy_yr linkflag /*edit this line to add any other desired nhis variables*/
```

```
save "meps\meps22nh", replace
```

```
describe
```

```
tab srvy_yr linkflag, missing
```

```
log close
```

6.0 Further Information

For any questions regarding the linkage file, please contact Julia Zhu at Julia.Zhu@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>.