Template for Requesting Verification Metrics for the Synthetic Linked NHIS-HUD-CMS Data

(Revised October 2026)

The National Center for Health Statistics (NCHS) Data Linkage Program has created a synthetic linked data set that includes variables from the 2018 National Health Interview Survey (NHIS) public-use survey data, the linked 2018 NHIS-Housing and Urban Development (HUD) restricted-use linked data files, the 2018 NHIS-Center for Medicare & Medicaid Services (CMS) Medicare restricted-use linked data files, and contextual information based on survey respondent Zip Code Tabulation Areas (ZCTAs) from the Agency for Healthcare Research and Quality (AHRQ) Community-Level Health (CLH) database. The restricted-use linked 2018 NHIS-HUD data files and the restricted-use linked 2018 NHIS-CMS Medicare data files continue to be available for researcher use via the NCHS Research Data Center Network.

To support the release of the synthetic linked NHIS-HUD-CMS data set, NCHS has developed a verification process to allow researchers to compare regression model results based on public-use synthetic linked data with model results based on original restricted-use linked data. Analysts at NCHS will run the user's specified model on both the synthetic and original restricted data and provide verification metrics back to the user including: 1) whether the signs of the estimate parameters agree, 2) whether the tests for statistical significance agree, 3) whether the synthetic coefficient estimate is contained within the confidence interval of the original coefficient estimate, and 4) the approximate percentage overlap of the confidence intervals. For more information on the synthetic linked NHIS-HUD-CMS data and the verification process, please review the NCHS Synthetic Linked 2018 NHIS-HUD-CMS Data: User Guide.

Models must contain linked variables (not just NHIS variables). Additionally, the verification process currently supports only logistic regression models. Interaction terms between two or more variables will not be supported. If running the specified model on the synthetic data returns certain types of errors, these verification metrics might not be displayed. For example, the synthetic model may report coefficients as "NA" if the corresponding predicator is linearly dependent with some function of other predictors or if the corresponding predictor has only one value for all records in the model estimation. In this case, no verification metrics will be reported. For more information on the synthetic linked NHIS-HUD-CMS data and the verification process, please review the NCHS Synthetic Linked 2018 NHIS-HUD-CMS Data: User Guide.

Due to limited staffing resources, the number of verification requests may be limited per data user. Users of the synthetic data should use the template below to specify the model of interest and request verification metrics. Requests should be sent to datalinkage@cdc.gov.

1. Outcome/Dependent variable

Please select your outcome variable:

For all adults (aged 18 and older) (choose one)

☐ NHIS Subjective Health Status: NHIS_HEALTH
☐ NHIS Disability Recode: NHIS_DISABILITY
☐ NHIS Flu Vaccine Recode: NHIS_FLUVAX
☐ NHIS Usual Place of Care Recode: NHIS_PLACEFORCARE
☐ NHIS Serious Psychological Distress Recode: NHIS_PSYCHDISTRESS_R
☐ HUD Housing Assistance Concurrent Recode: HUD_ASSISTCONCUR

For adults aged 65 and over who are linked to Medicare (UNIV_65P_MCARE=1) only:

☐ CMS Recoded Coverage Type: CMS_COVERAGE_TYPE
☐ CMS Vital Status: CMS_VITALSTATUS
☐ CMS Recoded Emergency Department Visits: CMS_EDVISITS_R
☐ CMS Recoded Inpatient Hospitalizations: CMS_IPVISITS_R
☐ CMS Recoded Payments (1): CMS_TOTAL_PAYMENTS_R1
☐ CMS Recoded Payments (2): CMS_TOTAL_PAYMENTS_R2

2. Predictor/Independent variables

Please select your predictor variable(s), organized below by data source. (Choose one or more.)

Source: NHIS Public-use File

☐ Participant Age: NHIS_AGE
☐ Age Recode (1): NHIS_AGE_R1
☐ Age Recode (2): NHIS_AGE_R2
Note: only one of the three age variables above may be selected in a single model.
☐ Usual Place of Care: NHIS_PLACEFORCARE
☐ Health Insurance Type: NHIS_INSURANCE
☐ Health Insurance Recode: NHIS_INSURANCE_R
☐ Disability Status: NHIS_DISABILITY
☐ Employment Status: NHIS_EMPLOYMENT
☐ Education Level: NHIS_EDUCATION
☐ Social Services: NHIS_SOCIALASSIST
☐ Flu Vaccine: NHIS_FLUVAX
☐ Participant Race/Hispanic origin: NHIS_RACE_ETH
☐ Rental Status: NHIS_HOUSEOWN
☐ Serious Psychological Distress: NHIS_PSYCHDISTRESS
☐ Serious Psychological Distress Recode: NHIS_PSYCHDISTRESS_R

☐ Number of Selected Chronic Conditions: NHIS_CHRONICCOND
☐ Chronic Condition Recode: NHIS_CHRONICCOND_R
☐ Subjective Health Status: NHIS_HEALTH
☐ Poverty Status: NHIS_POVERTY
☐ Poverty Status Recode: NHIS_POVERTY_R
☐ Pregnancy Status: NHIS_PREGNANT
☐ Marital Status: NHIS_MARITAL
☐ Participant Sex: NHIS_SEX
☐ Smoking Status: NHIS_SMOKING

Source: NHIS-HUD Linked Files

☐ Housing Assistance (Ever): HUD_LINKED
☐ Housing Assistance (Concurrent): HUD_ASSISTCONCUR
☐ Housing Assistance (After Interview): HUD_ASSISTAFTER
☐ Housing Assistance (Two Years Before Interview): HUD_ASSIST2YPRIOR
☐ Housing Assistance (Five Years Before Interview): HUD_ASSIST5YPRIOR

Source: NHIS-CMS Linked Files (These variables only apply to adults aged 65 and older who linked to Medicare and UNIV_65P_MCARE=1 or UNIV_MCARE_FFS=1 should be selected as the universe).

☐ Dual Enrollee: CMS_DUALENROLL
☐ Fee-for-Service Coverage (Months): CMS_FFS_MONTHS
☐ Medicare Advantage Coverage (Months): CMS_MA_MONTHS
☐ Recoded Coverage Type: CMS_COVERAGE_TYPE
☐ Any Medicare Enrollment: CMS_LINKED
☐ Number of Inpatient Hospitalizations: CMS_IPVISITS
☐ Recoded IP Visits: CMS_IPVISITS_R
☐ Number of Emergency Department Visits: CMS_EDVISITS
☐ Recoded ED Visits: CMS_EDVISITS_R
☐ Total Medicare Payments (Dollars): CMS_TOTALPAYMENTS
☐ Recoded Payments (1): CMS_TOTAL_PAYMENTS_R1
☐ Recoded Payments (2): CMS_TOTAL_PAYMENTS_R2
☐ Vital Status: CMS_VITALSTATUS

Source: AHRQ CLH Database

☐ ZCTA Internet Access: ACS_INTERNET
☐ ZCTA Income to Poverty Ratio: ACS_INCOME
☐ ZCTA Medicaid Coverage: ACS_MEDICAID
☐ ZCTA Uninsured: ACS_UNINSURED

3. Reference levels (optional):

Please clarify reference groups for selected variables (predictor/independent and/or outcome/dependent) in your model, if desired. If not specified, the lowest value will be used as the reference level.

4. Universe/Subpopulation for the model

Please select the universe or subpopulation who should be included in the model. (Choose one.)

☐ All adults aged 18 and older (no exclusions)
☐ Adults aged 65 and over who are linked to Medicare (UNIV_65P_MCARE=1)
☐ Adults aged 65 and over who are linked to Medicare and had Traditional / Fee-for-Service coverage for at least one month (UNIV_MCARE_FFS=1)
☐ Adults aged 18 and older who linked to one or more years of HUD administrative data (HUD_LINKED=1)
☐ Adults aged 18 and older with family income-to-poverty ratio less than 2 (UNIV_LT2XPOV=1)
Note: only available when HUD_ASSISTCONCUR is selected as the outcome variable.
☐ Adults aged 18 and older with family income-to-poverty ratio less than 2 AND housing status indicates rental (UNIV_RENTLT2XPOV=1)
Note: only available when HUD_ASSISTCONCUR is selected as the outcome variable.

Required Information

What is your e-mail address? _____