

Template for Requesting Verification Metrics for the Synthetic Linked NHCS-NDI Data

(October 2026)

The National Center for Health Statistics (NCHS) Data Linkage Program has created a public-use synthetic linked data set that includes variables from the [restricted-use 2016 National Hospital Care Survey \(NHCS\)](#), the [restricted-use linked 2016 NHCS – 2016-2017 National Death Index \(NDI\) data](#), and the [restricted-use 2016 NHCS Enhanced Opioid Identification](#) files. Each of these files continue to be available for researcher use via the NCHS Research Data Center Network.

To support the release of the synthetic linked NHCS-NDI 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 NHCS-NDI data and the verification process, please review the [NCHS Synthetic Linked 2016 NHCS-NDI Data: User Guide](#).

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 predictor 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. Please refer to the [NCHS Synthetic Linked 2016 NHCS-NDI Data: User Guide](#) for more information on the handling of verification metrics for anomalous regressions.

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, organized below by data source (choose one):

Source: Linked NHCS-NDI File

☐ Patient Final Mortality Status: MORTSTAT

Note: MORTSTAT cannot be selected as an outcome or predictor variable when it is used for universe/subpopulation selection.

☐ Underlying Cause of Death: Alzheimer's Disease: UCOD_ALZHEIMERS_DISEASE

☐ Underlying Cause of Death: Malignant Neoplasms: UCOD_CANCER

☐ Underlying Cause of Death: Chronic Lower Respiratory Diseases: UCOD_CHRONIC_LOWER_RESP_DISEASE

☐ Underlying Cause of Death: Diabetes Mellitus: UCOD_DIABETES

☐ Underlying Cause of Death: Diseases of the Heart: UCOD_HEART_DISEASE

☐ Underlying Cause of Death: Influenza and Pneumonia: UCOD_INFLUENZA_AND_PNEUMONIA

☐ Underlying Cause of Death: Nephritis, Nephrotic Syndrome and Nephrosis: UCOD_KIDNEY_DISEASE

☐ Underlying Cause of Death: Cerebrovascular Diseases: UCOD_STROKE

☐ Underlying Cause of Death: Unintentional Injuries: UCOD_UNINTENTIONAL_INJURIES

☐ Patient Death Within 30 Days of Last Encounter: DEATH_WITHIN_30DAYS

☐ Patient Death Within 60 Days of Last Encounter: DEATH_WITHIN_60DAYS

☐ Patient Death Within 90 Days of Last Encounter: DEATH_WITHIN_90DAYS

☐ Drug Overdose Deaths – Recode: DRUG_OVRDSE_DEATH_R

☐ Drug Overdose Deaths Involving Any Opioid - Recode: DRUG_OVRDSE_DEATH_ANYOPIOID_R

Source: NHCS Enhanced Opioid Identification File

☐ Total Number of Opioid-involved Hospital Encounters - Recode: TOT_OPIOID_ENC_R

2. Predictor/Independent variables

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

Source: NHCS Patient Encounter Records

☐ Patient Age: P_AGE_CD

☐ Patient Age - Recode: P_AGE_CD_R

Note: Only one of the two age variables above may be selected in a single model.

☐ Patient Sex: P_SEX_CD

☐ Imputed Patient Race/Ethnicity: P_RACE_ETHNCTY_CD

☐ Patient Residence Urban/Rural Status: P_URBRURAL_CD

☐ Patient Residence Urban/Rural Status - Recode: P_URBRURAL_CD_R

Note: Only one of the two urban/rural status variables above may be selected in a single model.

☐ Setting of Last Patient Encounter Record (Inpatient - IP or Emergency Department - ED): LAST_ENC_SETTING

☐ Month of Last Encounter (IP or ED): MONTH_LAST_ENC

☐ Chronic Obstructive Pulmonary Disease Diagnosis: COPD_IND

☐ Diabetes Diagnosis: DIABETES_IND

☐ Heart Disease Diagnosis: HEART_DISEASE_IND

☐ Mental Health Diagnosis: MENTAL_HEALTH_IND

☐ Sepsis Diagnosis: SEPSIS_IND

☐ Total Number of ED Encounter Records: TOT_ED_ENC

☐ Total Number of ED Encounter Records – Recode 1: TOT_ED_ENC_R1

☐ Total Number of ED Encounter Records – Recode 2: TOT_ED_ENC_R2

Note: Only one of the three total number of ED encounter records variables above may be selected in a single model.

- ☐ Total Number of IP Encounter Records: TOT_IP_ENC
- ☐ Total Number of IP Encounter Records – Recode 1: TOT_IP_ENC_R1
- ☐ Total Number of IP Encounter Records – Recode 2: TOT_IP_ENC_R2

Note: Only one of the three total number of IP encounter records variables above may be selected in a single model.
- ☐ Total Number of Patient Intensive Care Unit (ICU) Days: TOT_ICU_DAYS
- ☐ Total Number of Patient ICU Days – Recode 1: TOT_ICU_DAYS_R1
- ☐ Total Number of Patient ICU Days – Recode 2: TOT_ICU_DAYS_R2

Note: Only one of the three total number of patient ICU days variables above may be selected in a single model.
- ☐ ED Encounter Within 7 days of IP Discharge: ED_ENCOUNTER_POSTIP_IND
- ☐ Hospital Discharge Status: DSCHRG_STUS_CD
- ☐ Hospital Discharge Status - Recode: DSCHRG_STUS_CD_R

Note: Only one of the two hospital discharge status variables above may be selected in a single model.

Source: NHCS Enhanced Opioid Identification File

- ☐ Total Number of Opioid-involved Hospital Encounters: TOT_OPIOID_ENC
- ☐ Total Number of Opioid-involved Hospital Encounters - Recode: TOT_OPIOID_ENC_R

Note: Only one of the two total number of opioid-involved hospital encounters variables above may be selected in a single model.

Source: Linked NHCS-NDI File

- ☐ Patient Final Mortality Status: MORTSTAT

Note: MORTSTAT cannot be selected as an outcome or predictor variable when it is used for universe/subpopulation selection.
- ☐ Underlying Cause of Death: Nine (9) Leading Categories: UCOD_LEADING
- ☐ Underlying Cause of Death: Alzheimer's Disease: UCOD_ALZHEIMERS_DISEASE
- ☐ Underlying Cause of Death: Malignant Neoplasms: UCOD_CANCER
- ☐ Underlying Cause of Death: Chronic Lower Respiratory Diseases: UCOD_CHRONIC_LOWER_RESP_DISEASE
- ☐ Underlying Cause of Death: Diabetes Mellitus: UCOD_DIABETES
- ☐ Underlying Cause of Death: Diseases of the Heart: UCOD_HEART_DISEASE
- ☐ Underlying Cause of Death: Influenza and Pneumonia: UCOD_INFLUENZA_AND_PNEUMONIA
- ☐ Underlying Cause of Death: Nephritis, Nephrotic Syndrome and Nephrosis: UCOD_KIDNEY_DISEASE
- ☐ Underlying Cause of Death: Cerebrovascular Diseases: UCOD_STROKE
- ☐ Underlying Cause of Death: Unintentional Injuries: UCOD_UNINTENTIONAL_INJURIES

Note: Underlying cause of death variables cannot be selected as predictors if an underlying cause of death is selected as the outcome in a model.
- ☐ Number of Days Between Patient's Last Encounter and NDI Date of Death: TIME_TO_DEATH_CD
- ☐ Patient Death Within 30 Days of Last Encounter: DEATH_WITHIN_30DAYS
- ☐ Patient Death Within 60 Days of Last Encounter: DEATH_WITHIN_60DAYS
- ☐ Patient Death Within 90 Days of Last Encounter: DEATH_WITHIN_90DAYS

Note: Only one of the four time to death variables above may be selected in a single model.
- ☐ Drug Overdose Death Indic: DRUG_OVRDSE_DEATH
- ☐ Drug Overdose Deaths – Recode: DRUG_OVRDSE_DEATH_R
- ☐ Drug Overdose Deaths Involving Any Opioid: DRUG_OVRDSE_DEATH_ANYOPIOID
- ☐ Drug Overdose Deaths Involving Any Opioid - Recode: DRUG_OVRDSE_DEATH_ANYOPIOID_R

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 adult patients aged 18 and older (no exclusions)
- ☐ Only adult patients assumed to be deceased based on NDI linkage (MORTSTAT=1)

Note: MORTSTAT cannot be selected as an outcome or predictor variable when it is used for universe/subpopulation selection.

Required Information

What is your e-mail address? _____