

Synthetic Healthcare Data for Research (SyH-DR) Known Data Issue: Primary Diagnosis Imputed Flag

If you downloaded the SyH-DR data files before January 2024, the variable **Primary Diagnosis Code (ICD-10) Imputed Flag** (PRMRY_DX_IMPUTED) was assigned a nonmissing value (0 or 1) for all claim records, including cases with a missing value in the variable **Primary Diagnosis Code (ICD-10)** (PRMRY_DX_CD). The **Primary Diagnosis Code (ICD-10)** value is missing for any record with no primary or secondary diagnosis codes present for the claim. (ICD-10 stands for International Classification of Diseases, 10th Revision.)

As of the SyH-DR data file update in January 2024, the **Primary Diagnosis Code (ICD-10) Imputed Flag** value was set to missing for records with no diagnosis codes present on the claim. The update to the **Primary Diagnosis Code (ICD-10) Imputed Flag** variable affects records in the Medicaid and Commercial Inpatient and Outpatient files. The table below shows the number of records affected by the change for each of the payers, by file (SyH-DR).

Number of records changed to missing value for the Primary Diagnosis Code (ICD-10) Imputed Flag variable

SyH-DR Data File Name (.csv, .sas7bdat, and .dta)	Payer, Setting	Number of Records Changed to Missing Value*
syhdr_commercial_inpatient_2016	Commercial, Inpatient	156
syhdr_commercial_outpatient_2016	Commercial, Outpatient	3,889
syhdr_medicare_inpatient_2016	Medicare, Inpatient	0
syhdr_medicare_outpatient_2016	Medicare, Outpatient	0
syhdr_medicaid_inpatient_2016	Medicaid, Inpatient	11,100
syhdr_medicaid_outpatient_2016	Medicaid, Outpatient	440,967
Total		456,112

*The Primary Diagnosis Code (ICD-10) Imputed Flag variable contained no missing values before the January 2024 update.

If you downloaded SyH-DR data before January 2024 and would like a copy of the updated files, email the team at DataInnovations@ahrq.hhs.gov.

