Documentation for the Medicaid Enrollment File
February 2021

FILE NAME: Medicaid Enrollment File

MAXIMUM RECORD LENGTH (LRECL): 38

CANCER DIAGNOSIS YEARS: 2006-2013

SEER SUBMISSION DATE: November 2015

PROGRAMMER: Jennifer Stevens, Information Management Services, Inc.

CONTACT: SEER-Medicaid Technical Support, Information Management Services, Inc.
        Email: SEER-Medicaid@imsweb.com

SUMMARY:

Personal identifiers for all tumors diagnosed between 2006-2013, as reported to the SEER registries by November 2015, were linked to the CMS Medical Statistical Information System (MSIS) Medicaid Analytic Extract (MAX) Personal Summary (PS) Files (2006-2013*) from all 50 states and the District of Columbia using a deterministic process. Personal identifiers considered were Social Security Number (SSN), Health Insurance Claim (HIC), which is a unique Medicare number, date of birth and sex.

To be considered a match all persons had to have at least a complete match on SSN or HIC. The overall strength of the match was assessed. If there was an exact match on all digits of both the HIC and SSN, eight points were given. Four points were given if there was an exact match on HIC, but not the SSN or if there was a match on SSN and sex but not on the HIC. A match on sex was required with the SSN match because the SSN on the file could have been that of the spouse rather than the Medicaid recipient. Unlike the SSN, the HIC, by design, includes information about whether the individual is the husband or wife. An exact match on SSN without agreement on sex was given a lower value (three points). A match on sex was given one point. For date of birth, a hierarchical approach was used: exact match on (day/ month/year, three points), month and year (two points), or year only (one point). The points were summed; the maximum possible score was 12, the minimum was 3. Although the match score could vary over time because the yearly state MAX files could contain different information, the highest state-specific match score was assigned to each person for all years of data from that state.

The resulting Medicaid Enrollment File contains one record per person-state-year combination for all persons who were found to be Medicaid enrolled during at least one month between 2006-2013. Persons can have more than one record if they were found to be enrolled in Medicaid during more than one year and/or more than one state. Note there are some persons were found to have Medicaid enrollment in more than one state during the same year and even month.

*During the included years, states began transitioning from MSIS to the Transformed Medical Statistical Information System (TMSIS), meaning MAX-PS files were not available for all years from all states. Therefore, it is possible that some enrollment data may not have been captured. Please refer to this summary table of available MAX-PS files by state and year. Depending on the research question, investigators will need to decide how to handle persons who were assessed but were never found to be enrolled in Medicaid (i.e., personal identifiers were not found in any of the Medicaid state MAX-PS files).

For more information on the matching process, please refer to:

- Warren JL, Benner S, Stevens J, Enewold L, Huang B, Zhao L, Tilahun N, Bradley CJ. Development and Evaluation of a Process to Link Cancer Patients in the SEER Registries to National Medicaid Enrollment Data. J Natl Cancer Inst Monogr (2020).

Documentation for the Medicaid Enrollment File
February 2021

FILE DESCRIPTION

Field	SAS Variable Name	Position	Notes
Patient ID	PATIENT_ID	1-15	This field is used to uniquely identify a person. One person can have multiple primaries but the same Patient ID. See the sequence number for more information about the primary. This is a dummy number and is not the number used by the registry to identify the patient. This identifier is alphanumeric.
Medicaid state code	STATE_CD	16-17	2-character state abbreviation of Medicaid
Medicaid enrollment year	YEAR	18-21	4-digit year of Medicaid match
DOB match score (from highest yearly state specific scored match)	DOB_MATCH_SCORE_HI	22	0 = No match 1 = Year match 2 = Year/month match 3 = Exact match (year/month/day)
Gender match score (from highest yearly state specific scored match)	SEX_MATCH_SCORE_HI	23	0 = No match 1 = Match (M/M or F/F)
ID match score (from highest yearly state specific scored match)	ID_MATCH_SCORE_HI	24	3 = SSN Match 4 = HIC Match 8 = HIC and SSN Match
Match score	MATCH_SCORE_HI	25-26	Sum of DOB, Gender and ID match scores (value range: 3-12)
Medicaid Eligibility code – January through December	MEDICAID_ELIG_CD1 - MEDICAID_ELIG_CD12	27-38	0 = Not eligible 1 = Cash assistance (includes aged, blind) 2 = Eligible under 1931 3 = Medically needy 4 = Poverty 5 = NCCCP 6 = Other 7 = Foster care child 8 = 1115 expansion 9 = Unknown