Linking children from the National Health and Nutrition Examination Survey to Medicaid enrollment and claims data

Lisa B. Mirel[1], Cordell Golden[2], Gloria Wheatcroft[2], Alan E. Simon[2], and Kenneth C. Schoendorf[2]

[1]National Center for Health Statistics, Division of Health and Nutrition Examination Surveys, LMirel@cdc.gov
[2]National Center for Health Statistics, Office of Analysis and Epidemiology

Introduction

The National Health and Nutrition Examination Survey (NHANES) is a nationally representative survey that collects information on the health and nutritional status of adults and children in the United States, including information on health insurance coverage. Recently the 1999-2007 Centers for Medicare and Medicaid Services (CMS) Medicaid Analytic eXtract (MAX) files were linked to the 1999-2004 NHANES. This linkage was part of an interagency agreement between the National Center for Health Statistics (NCHS), CMS, and the U.S. Department of Health and Human Services (HHS) Assistant Secretary for Planning and Evaluation (ASPE), and the Social Security Administration (SSA). The linked data are available in the NCHS Research Data Center. The results of this study offer preliminary insights into opportunities to improve the collection of program participation information in a self-reported survey.

For this record linkage, NCHS obtained records for its survey respondents who had Medicaid records in the CMS data files. NHANES participants were defined as linkage eligible if sufficient identifiers were obtained at the interview and records from CMS could be sought. This paper will assess concordance between a participant's report of Medicaid enrollment at the time of their household interview in NHANES compared with the presence of a Medicaid enrollment record in the MAX files for children less than 18 years. We will compare the prevalence and characteristics of the concordant and discordant pairs. These results will be discussed in the context of challenges encountered in linkage and accurately collecting self-report of Medicaid enrollment in the survey.

Background on NHANES

NHANES is a continuous survey (1999-present) that includes an interview in the household followed by an examination in a Mobile Examination Center (MEC). NHANES is a representative cross-sectional sample of the U.S. civilian, non-institutionalized population that is selected using a complex, multistage probability design. Detailed descriptions of the NHANES sample design have been described elsewhere and only will be summarized here (Curtin et.al, 2012). Participants can be any age. An adult proxy responds to the majority of the questionnaire for children who are 16 years or younger.

The NHANES household interview contains a section which includes sample participant level interview data on health insurance coverage. One reference person in each family responds for each individual sample participant in the family. The reference person or responsible adult may or may not be a sample participant. Multiple families in the same household may be selected as participants in the study. However, this section of the family interview is asked for each individual sample participant and not for the family as a whole.

The initial set of questions in the health insurance section determined which sample participants had health insurance coverage. For those who had health insurance coverage, respondents were then asked "What kind of health insurance do you have?" The respondent was shown a hand card and asked to choose one from the list which included both Medicaid and Children's Health Insurance Program (CHIP). For 1999-2004 NHANES, the public-use available data reports Medicaid and CHIP together for the response. This was changed after 2004.

Background Medicaid/CHIP

Detailed information on Medicaid and the CHIP can be found on the CMS website (http://www.medicaid.gov/). Briefly, Medicaid is a needs based entitlement program administered by CMS in collaboration with each state to

provide health care coverage to some vulnerable populations in the United States, including low-income children, and the aged or disabled poor. The MAX files are research oriented Medicaid claims files. The eligibility for Medicaid varies by state. There is mandatory eligibility for pregnant women with a family income below 133% of the HHS poverty guidelines (this is sometimes referred to as the federal poverty level), children under 6 years of age with family income below 133% of the HHS poverty guidelines, and children aged 6-18 years of age with family income below 100% of the HHS poverty guidelines.

CHIP is a state based program offered to children and pregnant women whose incomes are too high to qualify for Medicaid eligibility. CHIP funds are capped with each state receiving a set allotment. The federal government matches state funding for CHIP. States receive higher federal funding matching rates for CHIP than Medicaid. Of particular note for this project, the MAX files are incomplete on CHIP because states have the option of not reporting stand-alone CHIP program data to the Medicaid Statistical Information System (MSIS) from which the MAX files are created.

Methods

For this analysis, we used the 1999-2004 NHANES data. NHANES participants were eligible for linkage if they did not refuse to provide a social security number (SSN) at the NHANES interview and provided sufficient personally identifiable information (PII) for linkage. SSNs for these participants were verified at SSA and then sent to CMS where data were extracted based on exact matches for SSN, gender and month and year of birth. Eligibility for linkage for a NHANES sample participant to obtain a CMS administrative record was independent of the questionnaire responses to the health insurance question and independent of a participant's Medicaid eligibility.

Our main research goal was to assess the reliability and accuracy of the NHANES Medicaid response as compared to the matched CMS records. Concordance was defined as having Medicaid/ CHIP reported as a health insurance source in NHANES and having a CMS administrative record within a month of the NHANES interview. The presence of a CMS administrative record within a month of the NHANES interview indicated active Medicaid enrollment status. The one month time period was chosen based on how the question was asked in NHANES which implied current enrollment. Records were also considered concordant if there was no report of Medicaid/ CHIP coverage in NHANES and no match to the CMS administrative records for Medicaid enrollment. Discordance was defined as a report of Medicaid/ CHIP in NHANES without CMS evidence of enrollment, or, conversely, no report of Medicaid/ CHIP in NHANES while actually being enrolled according to the MAX files.

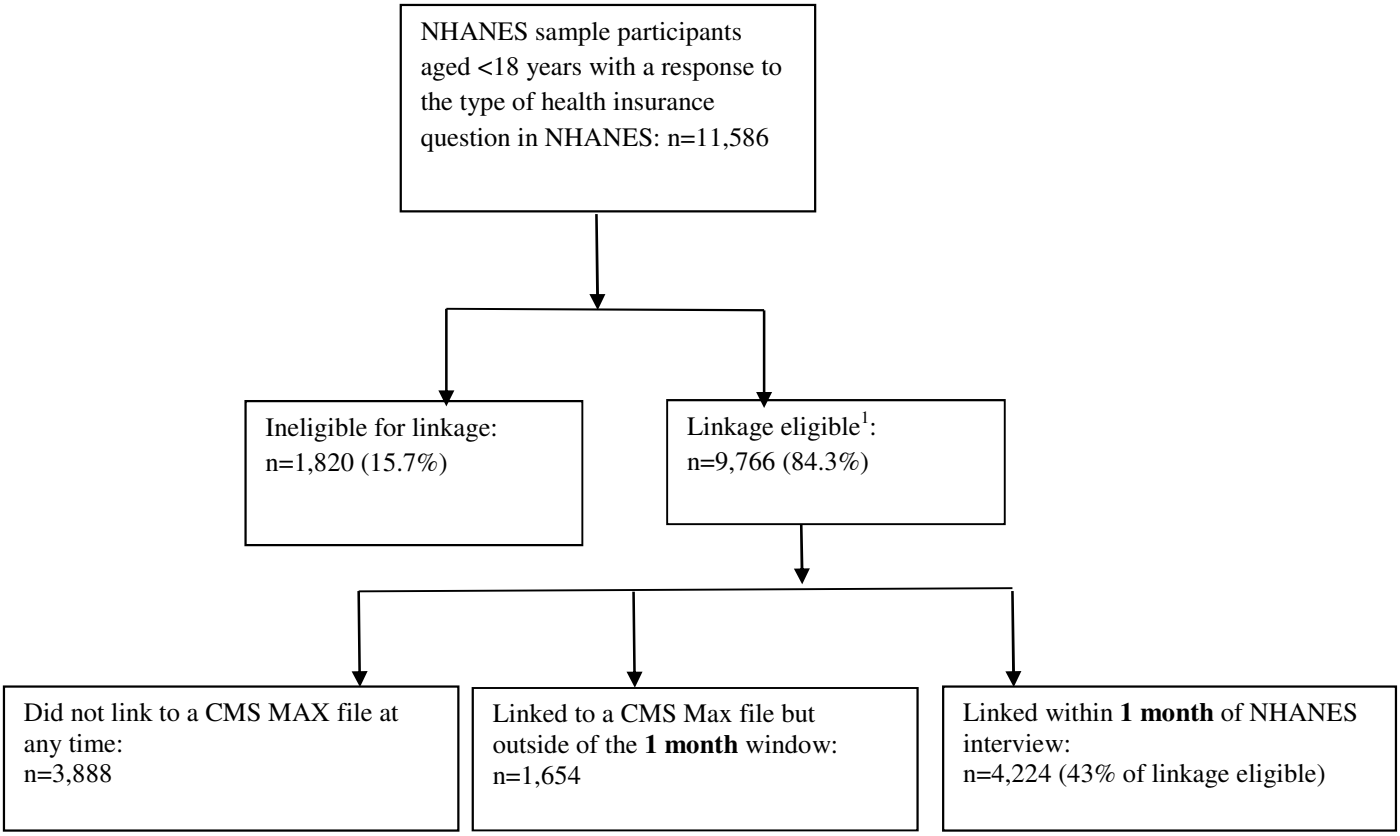
The first step of the analysis was to examine the overall analytic sample breakdown (Figure 1) and assess the overall matched results using weighted percentages, using the NHANES interview weights. We then looked at the percentage of concordance and discordance given the response to the Medicaid/CHIP question in NHANES. The denominator for these analyses included all linkage eligible sample participants. Socio-demographic characteristics were also examined, including the ratio of family income to the poverty index, household size, gender, age and race and Hispanic origin, for the concordant and discordant pairs.

A t-test was used to compare means and the Chi-square test of independence was used to assess statistically significant differences between demographic characteristics and concordance or discordance of Medicaid enrollment according to the MAX files and the interview reporting of type of health insurance coverage, using a two-sided p-value <0.05. Variances were estimated using Taylor Series Linearization, accounting for the complex design. All statistical analyses were performed using SAS (version 9.2, SAS Institute Inc., Cary, NC) and SAS-callable SUDAAN (version 10.0, Research Triangle Institute, Research Triangle Park, NC) software.

Results

From the 1999-2004 NHANES there were 11,586 participants aged less than 18 years who had a response to the type of health insurance question in NHANES (less than 1% were missing a response, n=65). About 84% of these participants were eligible for linkage (i.e., they did not refuse to provide a SSN at the NHANES interview, they provided sufficient PII for linkage and their SSN was verified at SSA). Of the 9,766 eligible for linkage, 4,242 participants matched to the Medicaid file within 1 month of the NHANES interview (Figure 1).

Figure 1. Analytic sample for 1999-2004 NHANES participants aged less than 18 years linked to CMS MAX file (note: all percentages in figure are unweighted)



[1] Linkage eligible means they did not refuse to provide a SSN at the NHANES interview, they provided sufficient personally identifiable information for linkage and their SSN was verified at SSA.

There are several different ways to assess the matched results. First we assessed the overall concordant and discordant weighted percentages for the linkage for all linkage eligible participants, n=9,766. The overall concordance, using weighted percentages was 89% (24%, n=3,101, said "Yes" in NHANES and linked within one month of the NHANES interview to the CMS MAX file and 65%, n=5,057, said "No" in NHANES and did not link to the CMS MAX file within 1 month of the NHANES interview or at any time).

Table 1. Unweighted sample sizes for linkage eligible NHANES sampled participants, comparing response to NHANES question about Medicaid/CHIP to linkage status to the CMS MAX file.

	Linked to the CMS MAX within 1 month of NHANES interview	Did not link to the CMS MAX file within 1 month or at any time	Total
"Yes" to question in NHANES about Medicaid/ CHIP health insurance coverage	3,101	485	3,586
"No" to question in NHANES about Medicaid/ CHIP health insurance coverage	1,123	5,057	6,180
Total	4,224	5,542	9,766

Of the respondents who said "No" to the question in NHANES, using weighted percentages,11% (1,123/6,180) linked within 1 month of the NHANES interview (making the agreement level of those who said "No" and were not linked 89%). Of the respondents who said "Yes" to the question in NHANES about Medicaid/CHIP health insurance coverage, using weighted percentages,12% (485/3,586) were not found in CMS MAX file (88% agreement among those who said "Yes" and linked within 1 month of the NHANES interview).

For those who reported Medicaid/CHIP coverage in NHANES but were not linked to the CMS MAX database they had a higher (p-value=0.002) mean income to poverty threshold ratio (mean=1.2, standard error=0.07), compared to those who reported Medicaid/CHIP and were linked (mean=1.0, standard error=0.03). The discordance group who responded "Yes" in NHANES but did not link to the CMS MAX file may be because these participants were receiving CHIP health insurance coverage and the MAX files are incomplete on CHIP data. The ratio of family income to poverty threshold of 1.2 for the discordant pair actually suggests they would be eligible in most states for CHIP, as income eligibility is slightly higher for CHIP than Medicaid. Table 2 compares other demographic characteristics of the two groups who reported "Yes" in NHANES. The other demographic characteristics between the two groups are similar which again suggests that perhaps the not linked group was receiving CHIP health insurance coverage. For those who reported "Yes" in NHANES there were no statistically significant differences based on concordance for a participants age or gender, however there were slight differences based on race and Hispanic origin (p-value= 0.04).

Table 2. Demographic characteristics of 1999-2004 NHANES participants aged 0-17[1] years who reported receiving Medicaid/CHIP health insurance coverage by concordance with the CMS MAX database

	Concordant "Yes" MAX file "Yes" NHANES n=3,101	Discordant "No" MAX file "Yes" NHANES n=485
Characteristic	Weighted Percent (SE)	Weighted Percent
Gender: Male	51.7 (1.0)	51.8 (4.1)
Age (years): <6	41.8 (1.6)	44.9 (3.8)
Race/ ethnicity		
non-Hispanic white	53.1 (3.9)	48.7 (5.4)
non-Hispanic black	30.3 (2.9)	26.4 (4.4)
Mexican American	16.6 (2.6)	25.0 (4.6)

[1] For participants aged 0-16 the responses were answered by a proxy

For those who did not report Medicaid/CHIP coverage in NHANES but were linked to the CMS MAX database, they had a lower (p-value <0.001) mean income to poverty threshold ratio (mean=1.3, standard error=0.05) compared to those who did not report Medicaid/CHIP and were not linked (mean= 3.2, standard error=0.06). It is possible that those who reported "No" in NHANES but were linked did not know they were enrolled in Medicaid/CHIP because they were enrolled in a managed care program through Medicaid. Table 3 compares other demographic characteristics of the two groups that reported "No" in NHANES. The characteristics are different by age and race and Hispanic origin (p-values <0.001). There were no statistically significant differences by gender. The demographic characteristics of the discordant pairs look more similar to those who reported receiving Medicaid/CHIP health insurance coverage in terms of age and race and Hispanic origin (see table 2).

Table 3. Demographic characteristics of 1999-2004 NHANES participants aged 0-17[1] years who did not report receiving Medicaid/CHIP health insurance coverage by concordance with the Medicaid Max database

	Concordant "No" MAX file "No" NHANES n=5,057	Discordant "Yes" MAX file "No" NHANES n=1,123
Characteristic	Weighted Percent (SE)	Weighted Percent
Gender: Male	50.9 (0.8)	55.5 (2.3)
Age (years): <6	29.4 (1.0)	36.7 (2.4)
Race/ ethnicity		
non-Hispanic white	82.0 (1.2)	50.3 (4.8)
non-Hispanic black	8.9 (0.9)	37.1 (4.9)
Mexican American	9.1 (1.1)	12.6 (2.2)

[1] For participants aged 0-16 the responses were answered by a proxy

Conclusions

There was some potential misreporting in the NHANES interview. However, some differences between the data from the two sources may be due to the lack of reporting of CHIP data in the CMS MAX files for those who reported "Yes" in NHANES but did not match to the Medicaid file. Our findings for those survey respondents reporting no Medicaid/ CHIP but linking to the CMS MAX data are similar to those results reported in the SNACC Medicaid undercount project using the Current Population Survey (SNACC is an acronym for the federal agencies conducting the project the University of Minnesota's State Health Access Data Assistance Center (SHADAC), NCHS, the Agency for Healthcare Research and Quality (AHRQ), ASPE, CMS, and the U.S. Census Bureau). More information on the SNACC project can be accessed at http://www.census.gov/did/www/snacc/index.html. Of those who reported "No" in NHANES and linked to Medicaid, 44% reported being covered by some other private insurance in the NHANES questionnaire and 54% reported they were covered by some other government insurance in the NHANES questionnaire.

In conclusion, 88% (3,101/3,586) of those who reported having Medicaid/CHIP health insurance coverage in NHANES linked to the Medicaid enrollment records within 1 month of the NHANES interview.

References

Curtin LR, Mohadjer L, Dohrmann S, et al. The National Health and Nutrition Examination Survey: Sample design, 1999–2006. National Center for Health Statistics. Vital Health Stat 2(155). 2012.