National Health Statistics Reports

Evaluating Survey Report of Social Security Disability Benefit Receipt Using Linked National Health Interview Survey and Social Security Administration Data

by Lisa B. Mirel, Cordell Golden, Jessica M. Keralis, Yeats Ye, Patricia C. Lloyd, and Julie D. Weeks

Abstract

Linking nationally representative population health survey data with Social Security Administration (SSA) disability program data provides a rich source of information on program recipients. Survey participant data from the 1998–2005 National Health Interview Survey (NHIS) were linked to SSA administrative records from 1997 through 2005. The goal of this study was to assess agreement between the actual benefit receipt based on the SSA administrative records and the survey report of benefit receipt in the linked NHIS and SSA file for the U.S. civilian noninstitutionalized population. This evaluation provides information on the expected accuracy of survey report of Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI) benefit receipt, including how participant characteristics may be associated with reporting misclassification. The results indicate that there is some underreporting of SSA disability benefit receipt based on the NHIS responses compared with the SSA administrative records. The analysis identified some differences between the concordant and discordant groups for selected characteristics, but there were no clear patterns among the different survey questions or the different survey participant characteristics.

Keywords: linked data • concordance • discordance • national survey

Introduction

The National Center for Health Statistics (NCHS) is the nation's principal health statistics agency, providing data to identify and address health issues. To maximize the scientific value of NCHS population-based surveys, NCHS has developed a data linkage program that links health survey data to vital and administrative data sources. Linked data files enable researchers to use longitudinal data from administrative databases or mortality data in combination with cross-sectional data to examine and provide more insight on factors that influence disability, health care utilization, morbidity, and mortality among different U.S. subpopulations compared with using one of these data sources alone.

One example is the linkage of nationally representative population health survey data with Social Security Administration (SSA) disability program data to provide a rich source of information about Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI) program recipients. Survey participant data from the 1998–2005 National Health Interview Survey (NHIS) were linked to SSA administrative records from 1997 through 2005. The resulting linked data provide key information on the demographics, health, and health care utilization of this growing population, which is not available in the administrative records alone. The linked data file also provides information on program participation and benefits history data from the administrative records, which is not available in the survey data alone. Research examining the accuracy of survey-reported data on receipt of benefits compared with SSA administrative records on receipt of benefits has established that survey-reported information on program participation and benefit receipt can differ from SSA administrative data (1–3). Much of the research has evaluated SSDI and SSI participation and benefit payment amounts using U.S. population surveys, including the Census Bureau's Survey of Income and Program Participation (SIPP) and the Current Population Survey

(CPS) linked to SSA data. The combined evidence from these analyses suggests that in some cases survey-reported data underreport Old-Age, Survivors, and Disability Insurance (OASDI) and SSI receipt (findings based primarily on SSA–CPS comparisons), while in other cases survey-reported data overestimate OASDI receipt, with varying results for SSI receipt (findings based primarily on SSA–SIPP comparisons) (1). In general, it appears that confusion among some survey participants exists as to which type of benefit they receive and the amount of income received (3).

Understanding the potential for underreporting is important for users of the survey data. The goal of this study was to assess agreement between the actual benefit receipt based on the administrative records and the survey-reported benefit receipt of SSDI and SSI in linked NHIS and SSA data for the U.S. civilian noninstitutionalized population. This evaluation provides information on the expected accuracy of the survey report in NHIS of SSDI and SSI benefit receipt, and it identifies participant characteristics that may be associated with reporting misclassification.

Data Sources

NHIS

NCHS has administered NHIS, a nationally representative, cross-sectional population health survey, continuously since 1957. NHIS serves as the principal source of information on the health of the U.S. civilian noninstitutionalized population. Households are selected through a probability sampling frame drawn from each state and the District of Columbia, based on information from the decennial census. The NHIS sample design has been described in more detail elsewhere (4). Within each sampled household, families are identified and complete a brief questionnaire to collect selected demographics and broad health measures. Additionally, one adult and one child are randomly selected to receive a more detailed health questionnaire. A parent or guardian answers the sample child questionnaire on behalf of the child (5). As part of the family questionnaire,

respondents are asked a series of questions on sources of family income for all family members. The questions ask about receipt of Social Security income, including whether the income was received as a disability insurance benefit (SSDI) and whether additional SSI income was received.

SSDI and SSI

SSA pays disability benefits through two programs: SSDI and SSI (6). The SSDI program is an earnings-eligible program providing benefits to disabled adults (aged 18–64) with a work history. SSDI beneficiaries who are under 65 become eligible to receive Medicare benefits before retirement age following a 24-month waiting period from their first SSDI benefit payment. For SSDI beneficiaries with selected rare diagnoses, this waiting period is waived. Additionally, disabled railroad workers under full retirement age qualify for Railroad Retirement (RR) disability benefits if they have at least 10 years of railroad service, or 5 years of service after 1995 when the employee's combined credits for work under Social Security and the RR program meet the eligibility requirements for Social Security disability benefits (7,8). Though the RR disability benefits are funded through the Railroad Retirement Board, this program shares several similarities with the SSDI program.

SSDI benefits are available for individuals unable to work because they have a condition expected to last at least 1 year or expected to result in death. Nonblind disabled individuals must satisfy two earnings criteria to be eligible: a) a recent work requirement based on age at time of disability, and b) a duration of work test to show that the individual contributed enough work earnings credits. Earnings credits are credits that individuals earn throughout their work history. When the individual turns 65, the SSDI benefits are thereafter referred to as retirement benefits rather than disability benefits. SSDI payments are based on wages received during working years. SSA's Master Beneficiary Record (MBR) file contains complete monthly SSDI benefit data.

Unlike SSDI, SSI is a federal income supplement program funded by tax revenues rather than an individual's earnings. The SSI program provides an income benefit to elderly, blind, or disabled individuals with limited income and resources (i.e., means tested). SSI benefits are available only to individuals whose income is below a certain level. The income benefits are designed to meet basic needs for food, clothing, and shelter, and the amount is not influenced by past earnings. In some states, meeting SSI eligibility requirements is satisfactory to meet Medicaid eligibility requirements. SSA's Supplemental Security Record (SSR) file contains complete monthly benefit data for all individuals receiving SSI benefits.

An individual may receive concurrent SSDI and SSI benefits. Because SSDI payments are based on wages received during working years, an SSDI payment may be low enough that the beneficiary meets the SSI eligibility criteria (6).

Linked NHIS and SSA files

The 1998–2005 NHIS data were linked to SSA data from 1997 through 2005 by the NCHS Data Linkage Program, following an established linkage algorithm (2). NHIS participants were eligible for linkage unless they refused to provide their Social Security Number (SSN) at the time of the NHIS interview, or were missing key personally identifiable information (PII) such as last name, first name, or date of birth (2). The NCHS Research Ethics Review Board approved the linkage of NHIS data with the SSA files. No additional approval was required by NCHS for this analysis.

Related questions in NHIS

NHIS participants were classified as having or not having reported SSDI or RR disability benefits based on responses to three nested NHIS questions:

Q1) "Did [person] receive income in [last calendar year] from Social Security or Railroad Retirement?"

NOTES: SSDI is Social Security Disability Insurance. RR is Railroad Retirement. Q1: "Did [person] receive income in [last calendar year] from Social Security or Railroad Retirement?" If "yes" to Q1, and the participant was under 65 at interview, the follow-up question Q2 asked, "Was [person's] Social Security or Railroad Retirement income received as a disability benefit?" If "yes" to Q2, the follow-up question Q3 asked, "Did [person] receive this benefit because [you are/he is/she is] disabled?" Missing includes nonresponse, don't know, or refusal to answer any of the three questions.
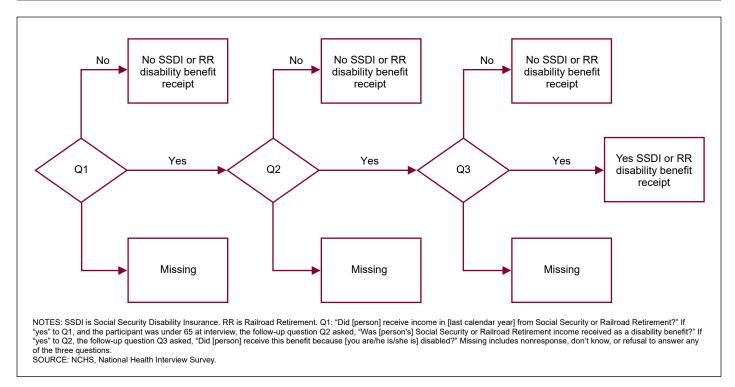SOURCE: NCHS, National Health Interview Survey.

Figure 1. Decision tree summarizing classification of National Health Interview Survey report of Social Security Disability Insurance or Railroad Retirement disability benefit receipt

Q2) If "yes" to Q1, and the participant was under 65 at interview, the follow-up question asked, "Was [person's] Social Security or Railroad Retirement income received as a disability benefit?"

Q3) If "yes," the follow-up question asked, "Did [person] receive this benefit because [you are/he is/she is] disabled?"

A survey participant was classified as having reported SSDI or RR disability benefits for their disability in the last calendar year if the response to all three questions was "yes." The analysis was limited to all three questions being answered in the affirmative to ensure that the benefit receipt was for the participant's disability and not a spouse or other family member's disability. If the response to any of the three questions was "no," the participant was classified with no report of SSDI or RR benefits. A participant was missing information for the report of SSDI benefits if they were missing a response to the first question, or missing a response to the second or third question when a response was expected. Missing includes nonresponse, don't know, or refusal to answer any of the three questions. Because it is not

possible to distinguish between SSDI and RR disability benefits from these questions, both programs were combined for this analysis (Figure 1).

Report of SSI disability benefits was derived from two nested questions:

Q1) "Did [person] receive Supplemental Security Income (SSI)?"

Q2) If "yes," the follow-up question asked, "Did [person] receive SSI because [you have/he has/she has] a disability?"

A survey participant was classified as having a survey report of SSI disability benefits for their disability in the last calendar year if the response to the two questions was "yes"; otherwise, the participant was classified as not having a survey report of the benefit. Similar to the SSDI or RR questions, using affirmative responses to both questions to classify receipt of benefit was done to ensure that the receipt of benefit was because of the participant's disability and not that of another family member. A survey participant was missing information for a report of SSI disability benefits if they were missing a response to the first question, or missing a response to the second question when an answer was expected (Figure 2). As noted previously,

missing included those with nonresponse to the questions, response of "don't know," or a refusal to answer either question.

A survey participant was classified as having a report of any SSA disability benefit receipt in the last calendar year if they met either of the SSDI or RR, or SSI criteria as defined in Figures 1 and 2, respectively.

SSA record of disability benefit receipt

Monthly variables from the SSA MBR file were used to determine SSA record of SSDI or RR disability benefit receipt and were used to identify disabled SSDI and RR beneficiaries in current pay status. Monthly variables from the SSR files were used to determine SSA record of SSI disability benefit receipt (2).

If an NHIS participant had an SSA record indicating receipt of SSDI or SSI disability benefits in at least 1 month during the calendar year before their NHIS interview, they were classified as receiving SSDI or RR, or SSI disability benefits. An NHIS participant with no SSA record of benefit receipt during the calendar year before their NHIS interview was classified as not receiving SSDI or RR, or SSI disability benefits

NOTES: SSI is Supplemental Security Income. Q1: "Did [person] receive Supplemental Security Income (SSI)?" If "yes" to Q1, the follow-up question Q2 asked, "Did [person] receive SSI because [you have/he has/she has] a disability?" Missing includes nonresponse, don't know, or refusal to answer either question.
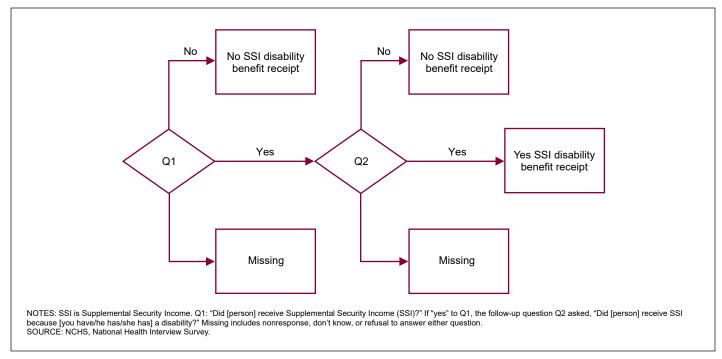SOURCE: NCHS, National Health Interview Survey.

Figure 2. Decision tree summarizing classification of National Health Interview Survey report of Supplemental Security Income disability benefit receipt

based on the SSA administrative data. The assessment of receipt of benefits for each program, SSDI or RR and SSI, was conducted separately.

Analytic sample

The analytic sample was restricted to 1998–2005 NHIS participants aged 18–64 who were eligible for linkage to the SSA administrative records and had complete NHIS data on disability benefit receipt (Figure 3). Of the 468,797 adults aged 18–64 interviewed in the 1998–2005 NHIS, 184,083 (39%) were excluded due to ineligibility for linkage. An additional 2,789 NHIS participants were excluded due to a missing response for NHIS disability benefit information, a response of don't know, or refused, leaving 281,925 participants in the analytic sample. The analysis was restricted to adults aged 18–64 because SSDI disability benefits are converted to retirement benefits at age 65, and the follow-up questions were asked only of participants under 65.

Participant characteristics at NHIS interview

Demographic characteristics of the study sample included age (18–39, 40–64), sex (male, female), and race and ethnicity (non-Hispanic white; non-Hispanic black; Hispanic; non-Hispanic, all other races). Socioeconomic status was described using poverty status as measured by the percentage of the federal poverty threshold (less than 100%, 100%–199%, 200% or more) from the NHIS imputed income file, educational attainment (no high school diploma, high school diploma or GED, more than a high school education), employment in the last week (yes, no), marital status (married, not married), and reported assessment of health status (excellent or very good, good, fair or poor).

Analytic Methods

Concordance and discordance of SSA and NHIS data was determined by comparing whether there was any benefit receipt (SSDI or RR, or SSI) from SSA to whether there was any report of SSDI or RR, or SSI benefit receipt in the survey. For each of the three analyses described below, participants were classified into

groups based on congruence of the SSA reports of disability benefit receipt and their responses to the NHIS questions related to the receipt of disability benefits.

1. Concordant: Yes–SSA record of benefit receipt and Yes–NHIS report of benefit receipt.
2. Discordant: Yes–SSA record of benefit receipt and No–No NHIS report of benefit receipt.
3. Discordant: No–No SSA record of benefit receipt and Yes–NHIS report of benefit receipt.
4. Concordant: No–No SSA record of benefit receipt and No–No NHIS report of benefit receipt.

The first analysis assessed agreement of any SSA disability benefit receipt between the SSA administrative record and the survey report of benefits, and included all linkage-eligible participants with nonmissing information on the survey report of any SSA disability benefit receipt (n = 281,925; Table 1). The second analysis assessed agreement of SSDI or RR disability benefit receipt between an SSA administrative record and the report of SSDI or RR disability benefit receipt in the survey, and included all linkage-eligible participants who had nonmissing information on the report of

SSDI or RR disability benefit receipt ($n = 281,668$; Table 2). The last analysis assessed agreement of SSI disability benefit receipt in the SSA record and the survey report, and included all linkage-eligible participants who had nonmissing information on the report of SSI disability benefit receipt ($n = 280,968$; Table 3).

Cohen's kappa statistic was used to measure agreement of receipt of disability benefits between the SSA administrative record and the survey report. The theoretical range of the kappa statistic is 0 for no agreement and 1 for complete agreement. A kappa statistic with a value of 0.60–0.79 is considered to have moderate agreement, 0.80–0.90 is considered strong, and above 0.90 is considered almost perfect (9).

A sensitivity analysis was conducted to evaluate results when the analytic sample was restricted to the participants who answered for themselves (i.e., proxy responses, which were responses provided for a participant by the NHIS family respondent, were excluded). This limited the sample size for the assessment of any SSA disability benefit receipt to 135,008; for the SSDI or RR disability comparison to 134,905; and for the SSI disability benefit receipt to 134,334.

Additionally, under the assumption that the SSA record of benefit receipt is the gold standard, the following statistics (10) were used to summarize how the SSA record compared with the NHIS report:

- Sensitivity: The proportion of participants with an SSA record of

a disability benefit who also had a survey report of that disability benefit.

- Specificity: The proportion of participants with no SSA disability benefit receipt who also had no survey report of that disability benefit.

- Positive predictive value: The proportion of participants with an SSA record of a disability benefit receipt among those with a survey report of a disability benefit receipt.

- Negative predictive value: The proportion of participants with no SSA record of a disability benefit receipt among those with no survey report of a disability benefit receipt.

To examine the potential differences between the groups, demographic and

NOTES: NHIS is National Health Interview Survey. SSA is Social Security Administration. SSDI is Social Security Disability Insurance. RR is Railroad Retirement. SSI is Supplemental Security Income.
SOURCE: NCHS, Linked National Health Interview Survey and Social Security Administration data.
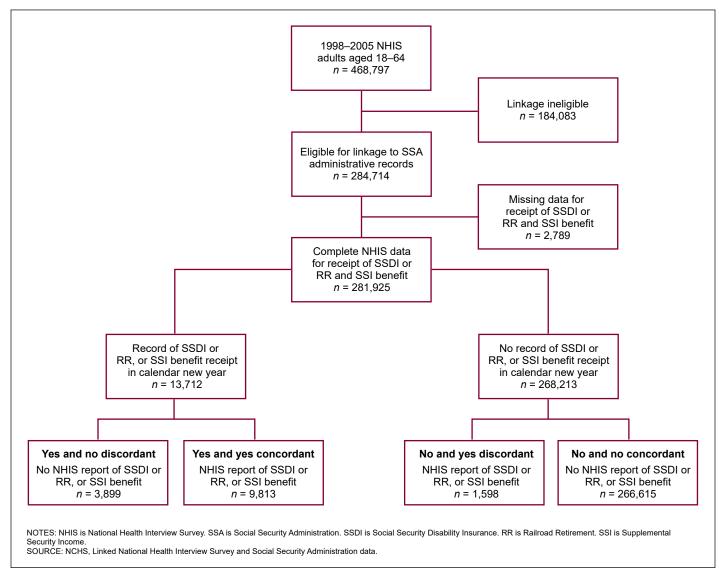
Figure 3. Analytic sample: 1998–2005 National Health Interview Survey participants aged 18–64 linked to any Social Security Administration file

socioeconomic variables for the Yes–SSA and No–NHIS pairs were compared with the Yes–SSA and Yes–NHIS pairs. Similarly, the discordant No–SSA and Yes–NHIS pairs were compared with the No–SSA and No–NHIS pairs.

Differences were evaluated with two-sided Wald tests at the 0.05 significance level. NHIS sample weights, adjusted for linkage eligibility using model-based calibration and poststratification procedures, were used in all analyses (2). Unless otherwise stated, the percentages and kappa statistics presented in this report were weighted based on the adjusted weights for linkage eligibility. The sensitivity, specificity, positive predictive value, and negative predicative value are based on unweighted counts. The models used to calculate the adjustment factors and poststratification totals for the weights included race and ethnicity, age, and sex. Variances were estimated in SUDAAN (release 10.0) to account for the complex design of NHIS.

Results

Any SSA disability benefit: Assessment of agreement

Of the 281,925 linkage-eligible participants with nonmissing information about report of SSDI or RR, or SSI disability benefits, data for 9,813 (3.5%) participants were concordant Yes–SSA and Yes–NHIS; 3,899 (1.4%) were discordant Yes–SSA and No–NHIS; 1,598 (0.6%) were discordant No–SSA and Yes–NHIS; and 266,615 (94.6%) were concordant No–SSA and No–NHIS (Table 1). Overall, the concordance was 98.1% (Yes–SSA and Yes–NHIS and No–SSA and No–NHIS). Agreement, as measured by the kappa statistic, was 0.77 (standard error [SE] = 0.0033) between the SSA administrative record of any SSA benefit receipt and the survey report of any receipt.

The sensitivity was 71.6% (Table 1). The positive predictive value was 86.0% (Table 1). The specificity and negative predictive value were 99.4% and 98.6%, respectively.

Restricting the analysis to the 135,008 participants who answered the NHIS income questions on disability benefits for themselves, the kappa statistic for report of any SSA disability benefit increased slightly, from 0.77 (SE = 0.0033) to 0.80 (SE = 0.0040, data not shown).

Discordant Yes–SSA and No–NHIS compared with concordant Yes–SSA and Yes–NHIS for any SSA benefit

There were no statistically significant differences between discordant Yes–SSA and No–NHIS and concordant Yes–SSA and Yes–NHIS groups by age, race and ethnicity, or sex. The discordant Yes–SSA and No–NHIS participants were more likely to be 200% or more above the poverty threshold, have a high school diploma or GED, be employed, be married, and report excellent or very good health compared with those who were linked and reported SSA disability benefit receipt in the survey (Table 4).

Discordant No–SSA and Yes–NHIS compared with concordant No–SSA and No–NHIS for any SSA benefit

For those who did not link to an SSA disability benefit receipt but did report the benefit in the survey, there were statistically significant differences for all variables when compared with those who did not link and did not report benefit receipt in the survey (Table 4). The No–SSA and Yes–NHIS participants were more likely to be aged 40–64, male, non-Hispanic black, below 100% of the poverty threshold, have no high school diploma or GED, not be employed in the last week, and report fair or poor health status compared with those who did not link and did not report benefit receipt in the survey.

SSDI or RR disability benefits: Assessment of agreement

Of the 281,668 linkage-eligible participants with nonmissing information from the survey for SSDI or RR disability benefits, 5,471 (1.9%) participants were concordant Yes–SSA and Yes–NHIS; 3,777 (1.3%) were discordant Yes–SSA and No–NHIS; 1,549 (0.5%) were

discordant No–SSA and Yes–NHIS; and 270,871 (96.2%) were concordant No–SSA and No–NHIS (Table 2). Similar to any SSA disability benefit receipt, the overall concordance (Yes–SSA and Yes–NHIS and No–SSA and No–NHIS) for those who reported receipt of SSDI or RR benefits was 98.1%. The kappa statistic for agreement between the SSA record of receipt of SSDI or RR benefits and the survey report was 0.67 (SE = 0.0049).

The sensitivity for the SSA record of benefit receipt and the survey report of SSDI or RR disability receipt was 59.2%. The positive predictive value for SSDI or RR benefit receipt was 77.9% (Table 2). The specificity and negative predictive value of the NHIS report for SSDI or RR disability benefits were 99.4% and 98.6%, respectively (Table 2).

When restricted to household participants who answered the question for themselves (n = 134,905, 48% unweighted), the kappa statistic to evaluate concordance for SSDI or RR disability benefits in the administrative and survey data increased to 0.72 (SE = 0.0055, data not shown).

Discordant Yes–SSA and No–NHIS compared with concordant Yes–SSA and Yes–NHIS for SSDI or RR

There were no statistically significant differences between discordant Yes–SSA and No–NHIS and concordant Yes–SSA and Yes–NHIS adults by marital status. There were statistically significant differences based on age, sex, race and ethnicity, poverty status, education, employment, and health status between those who linked to SSA but did not report SSDI or RR benefit receipt in the survey compared with those who were linked and reported SSDI or RR benefit receipt in the survey (Table 5). The Yes–SSA and No–NHIS group tended to have a higher percentage of participants who were aged 18–39, female, Hispanic or non-Hispanic black, below 100% of the poverty threshold, employed in the last week, and reported good, very good, or excellent health status compared with the Yes–SSA and Yes–NHIS group. The Yes–SSA and No–NHIS group had a lower percentage for those with more than a high school education compared with the Yes–SSA and Yes–NHIS group.

Discordant No–SSA and Yes–NHIS compared with concordant No–SSA and No–NHIS for SSDI or RR

For those who did not link but reported SSDI or RR benefit receipt in the survey, there were statistically significant differences for all variables compared with the No–SSA and No–NHIS group (Table 5). The No–SSA and Yes–NHIS group tended to have higher percentages of participants who were aged 40–64, male, non-Hispanic black, below 100% or between 100%–199% of the poverty threshold, had no high school diploma or a high school diploma or GED, were not employed in the last week, not married, and reported good, fair, or poor health compared with the No–SSA and No–NHIS group.

SSI disability benefits: Assessment of agreement

Of the 280,968 linkage-eligible participants with nonmissing information from the survey of SSI disability benefits, 4,328 (1.5%) participants were concordant Yes–SSA and Yes–NHIS; 1,914 (0.7%) were discordant Yes–SSA and No–NHIS; 1,391 (0.5%) were discordant No–SSA and Yes–NHIS; and 273,335 (97.3%) were concordant No–SSA and No–NHIS. The overall concordance (Yes–SSA and Yes–NHIS and No–SSA and No–NHIS) for this group was 98.8%. The kappa statistic for agreement between the NHIS report and SSA record of receipt of SSI benefits was 0.72 (SE = 0.0052).

The sensitivity of the SSA record and the survey report of SSI report was 69.3% (Table 3). The positive predictive value for SSI was 75.7% (Table 3). The specificity and the negative predictive value for SSI were 99.5% and 99.3%, respectively (Table 3).

When restricted to nonproxy family participants who answered the question for themselves ($n = 134,334$, 48% unweighted), the kappa statistic to evaluate concordance for SSI disability benefit in the administrative data and the survey increased to 0.75 (SE = 0.0065, data not shown).

Discordant Yes–SSA and No–NHIS compared with concordant Yes–SSA and Yes–NHIS for SSI

There was no statistically significant difference between those who linked to an SSA record of receipt of SSI benefit but did not report the benefit in the survey by age compared with the Yes–SSA and Yes–NHIS. However, for all other variables compared there were statistically significant differences between those who linked to an SSI benefit and did not report receiving the benefit in the survey compared with those who linked and reported receiving SSI disability benefits in the survey (Table 6). The Yes–SSA and No–NHIS group tended to have a higher percentage of participants who were male, non-Hispanic black, at 200% or higher than the poverty threshold, had a high school diploma or GED, more than a high school education, were employed in the last week, were married, and reported very good or excellent health status compared with the Yes–SSA and Yes–NHIS group.

Discordant No–SSA and Yes–NHIS compared with concordant No–SSA and No–NHIS for SSI

For those who did not link to an SSI disability benefit but reported receipt of an SSI benefit in the survey, there were no statistically significant differences by sex when compared with those in the No–SSA and No–NHIS group for SSI, but there were statistically significant differences for all other variables compared for the two groups (Table 6). The No–SSA and Yes–NHIS group tended to have a higher percentage of participants who were aged 40–64, non-Hispanic black, below 100% or between 100%–199% of the poverty threshold, had no high school diploma or a high school diploma or GED, were not employed in the last week, were not married, and reported good or fair or poor health status compared with the No–SSA and No–NHIS group.

Discussion

The overall agreement between the SSA administrative data and the survey reports of any SSA disability benefit receipt was moderate based on the kappa statistic. The results indicated some underreporting of SSA disability benefit receipt based on NHIS compared with SSA administrative records. The sensitivity of the questions on disability receipt in NHIS was 71.6%. This means that of the participants with an SSA record of receiving disability benefits, 28.4% were not correctly identified by their responses to NHIS questions. The basis for this type of discordance is unknown; however, the discordant Yes–SSA and No–NHIS group did not differ from the concordant Yes–SSA and Yes–NHIS group for any SSA benefit with respect to age, race and ethnicity, and sex. For other analyses, there were some differences between the concordant–discordant groups for selected characteristics, but there were no clear patterns among the type of SSA benefit and the differences for specific characteristics.

The overall agreement for SSDI or RR disability benefit receipt and SSA record of SSDI or RR was 98.1%, with a moderate kappa statistic (0.67). However, 40.8% (SE = 0.59) of participants who received SSDI disability benefits were not identified using their responses to the NHIS questions on SSDI or RR.

The NHIS questions nearly always identified individuals who were not receiving SSI disability benefits (specificity of 99.5%). The sensitivity of the questions on SSI receipt in NHIS was 69.3%. This means that of the participants with an SSA record of receiving SSI benefits, 30.7% were not correctly identified by their responses to NHIS questions. This percentage is higher than findings from Huynh et al., where approximately 15% of SSI-only recipients (based on the administrative data) reported OASDI-only benefits or no benefit at all in SIPP (3). This discrepancy may be attributed to differences in recall ability; SIPP asks about the previous 4 months, while NHIS asks about the previous calendar year. Nearly one-quarter of adults with an NHIS report of an SSI benefit in the

last calendar year did not link to an SSI record.

This analysis has several limitations. For example, NHIS participants eligible for linkage to the SSA records are a self-selected subset of the initial survey participants and may differ from those who were not eligible for linkage. Further, the percentage of NHIS participants ineligible for record linkage (based on SSN response or missing personal identifiers) increased during the study period from approximately 30% in 1998 to upwards of 56% in 2005 (11). The sample weights were adjusted for linkage ineligibility, which adjusts the sample weights using variables that influence the probability of linkage eligibility while maintaining population totals, but this may not fully adjust for selection bias.

Additionally, there may be errors in the SSA administrative data and errors in the linkage process. Both of these sources of error may have resulted in misclassification of SSA disability benefit receipt for NHIS participants. Based on the SSA records, there were no instances where a linked sample beneficiary received RR disability payments within the last calendar year. However, the NHIS question could not distinguish RR beneficiaries from SSDI beneficiaries. Therefore, there may have been errors in how survey report of benefits was characterized.

Furthermore, the authors took a conservative approach to classify NHIS report of SSDI or RR and SSI, requiring three questions about SSDI disability receipt and two questions about SSI disability to be answered affirmatively. This approach did not appear to affect the results. As a sensitivity analysis, a less conservative approach was also assessed. The less conservative approach limited the NHIS report of SSDI or RR benefit receipt to the first two questions and SSI to the first question. The results were almost identical. For any SSA benefit receipt, based on the less conservative classification, the kappa was 0.78 (SE = 0.0032). For SSDI or RR benefit receipt, it was 0.67 (SE = 0.0048) and for SSI benefit receipt, it was 0.73 (SE = 0.0050) (data not shown).

The NHIS benefit questions are asked of the family respondent, and it is possible for this person to know that the referent individual is receiving a benefit from SSA, but not know the type of benefit or if the benefit is due to the referent participant's own disability. The sensitivity analysis that was restricted to nonproxy responses to the disability benefit questions found minimal differences in overall concordance for report of SSI disability benefits, which may suggest that the question was answered with the same level of concordance whether proxy or nonproxy participants responded to the NHIS question about SSA benefits.

The linked NHIS and SSA data provide an opportunity to evaluate the expected accuracy of the data collected in NHIS relating to the receipt of SSA disability benefits through the SSDI and SSI programs. In lieu of an updated NHIS and SSA linkage, results from this analysis support research using only the survey questions on receipt of SSA benefit to examine characteristics of SSA beneficiaries for more recent years of NHIS where linked data are unavailable, despite some evidence of underreporting. However, future linkages will be useful to continuously monitor the accuracy of the data reported in the survey and may help support future questionnaire designs.

References

1. Davies PS, Fisher TL. Measurement issues associated with using survey data matched with administrative data from the Social Security Administration. Soc Secur Bull 69(2):1–12. 2009.
2. Golden C, Driscoll AK, Simon AE, Judson DH, Miller EA, Parker JD. Linkage of NCHS population health surveys to administrative records from Social Security Administration and Centers for Medicare & Medicaid Services. National Center for Health Statistics. Vital Health Stat 1(58). 2015.
3. Huynh M, Rupp K, Sears J. The assessment of Survey of Income and Program Participation (SIPP) benefit data using longitudinal administrative records. Washington, DC: U.S. Census Bureau. 2002.
4. Botman SL, Moore TF, Moriarity CL, Parsons VL. Design and estimation for the National Health Interview Survey, 1995–2004. National Center for Health Statistics. Vital Health Stat 2(130). 2000.
5. Black LI, Zablotsky B. Chronic school absenteeism among children with selected developmental disabilities: National Health Interview Survey, 2014–2016. National Health Statistics Reports; no 118. Hyattsville, MD: National Center for Health Statistics. 2018.
6. Social Security Administration Office of Retirement and Disability Policy. 2018 Red Book: A summary guide to employment supports for persons with disabilities under the Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI) programs. Washington, DC: Social Security Administration. 2018.
7. Whitman K. An overview of the Railroad Retirement program. Soc Secur Bull 68(2):41–51. 2008.
8. Social Security Administration. Social Security credits. Washington, DC. Available from: https://www.ssa.gov/planners/credits.html.
9. McHugh ML. Interrater reliability: The kappa statistic. Biochem Med (Zagreb) 22(3):276–82. 2012.
10. Gordis L. Epidemiology. Philadelphia, PA: W.B. Saunders. 1996.
11. Miller DM, Gindi RM, Parker JD. Trends in record linkage refusal rates: Characteristics of National Health Interview Survey participants who refuse record linkage. Presented at the 2011 American Statistical Association Joint Statistical Meeting. Miami, FL. 2011.

Table 1. Concordance and discordance of National Health Interview Survey report and Social Security Administration record of any Social Security Administration disability benefit receipt

NHIS report of any SSA disability benefit during last calendar year	SSA record of any SSA disability benefit receipt in last calendar year		
	Yes	No	Total
Yes	9,813	1,598	11,411
No.	3,899	266,615	270,514
Total	13,712	268,213	281,925

Kappa statistic: 0.77 (SE = 0.0033)
Sensitivity: 71.6% (9,813/13,712)
Specificity: 99.4% (266,615/268,213)
Positive predictive value: 86.0% (9,813/11,411)
Negative predictive value: 98.6% (266,615/270,514)

NOTES: NHIS is National Health Interview Survey. SSA is Social Security Administration. SE is standard error.

SOURCE: NCHS, Linked National Health Interview Survey and Social Security Administration data.

Table 2. Concordance and discordance of National Health Interview Survey report and Social Security Administration record of report of Social Security Disability Insurance or Railroad Retirement disability benefit receipt

NHIS report of SSDI or RR disability benefit receipt during last calendar year	SSA record of SSDI or RR disability benefit receipt during last calendar year		
	Yes	No	Total
Yes	5,471	1,549	7,020
No.	3,777	270,871	274,648
Total	9,248	272,420	281,668

Kappa statistic: 0.67 (SE = 0.0049)
Sensitivity: 59.2% (5,471/9,248)
Specificity: 99.4% (270,871/272,420)
Positive predictive value: 77.9% (5,471/7,020)
Negative predictive value: 98.6% (270,871/274,648)

NOTES: NHIS is National Health Interview Survey. SSDI is Social Security Disability Insurance. RR is Railroad Retirement. SSA is Social Security Administration. SE is standard error.

SOURCE: NCHS, Linked National Health Interview Survey and Social Security Administration data.

Table 3. Concordance and discordance of National Health Interview Survey report and Social Security Administration record for the report of Supplemental Security Income benefit receipt

NHIS report of SSI benefit receipt during the last calendar year	SSA record of SSI benefit receipt during the last calendar year		
	Yes	No	Total
Yes	4,328	1,391	5,719
No.	1,914	273,335	275,249
Total	6,242	274,726	280,968

Kappa statistic: 0.72 (SE = 0.0052)
Sensitivity: 69.3% (4,328/6,242)
Specificity: 99.5% (273,335/274,726)
Positive predictive value: 75.7% (4,328/5,719)
Negative predictive value: 99.3% (273,335/275,249)

NOTES: NHIS is National Health Interview Survey. SSI is Supplemental Security Income. SSA is the Social Security Administration. SE is standard error.

SOURCE: NCHS, Linked National Health Interview Survey and Social Security Administration data.

Table 4. Weighted percentages and standard errors for selected characteristics for linkage-eligible National Health Interview Survey participants aged 18–64, by concordant–discordant groups (linked National Health Interview Survey and Social Security Administration data), 1998–2005

Characteristic	Yes–SSA and Yes-NHIS		Yes–SSA and No–NHIS		No–SSA and Yes–NHIS		No–SSA and No–NHIS	
	n	Percent (SE)	n	Percent (SE)	n	Percent (SE)	n	Percent (SE)
Total	9,813	100.0	3,899	100.0	1,598	100.0	266,615	100.0
Age group								
18–39	2,334	23.7 (0.48)	912	23.2 (0.73)	420	[2]25.7 (1.22)	140,751	51.4 (0.14)
40–64	7,479	76.3 (0.48)	2,987	76.8 (0.73)	1,178	[2]74.3 (1.22)	125,864	48.6 (0.14)
Sex								
Male	4,687	48.4 (0.54)	1,849	48.2 (0.85)	843	[2]53.8 (1.37)	130,632	49.2 (0.08)
Female	5,126	51.6 (0.54)	2,050	51.8 (0.85)	755	[2]46.2 (1.37)	135,983	50.8 (0.08)
Race and ethnicity								
Hispanic	1,639	8.9 (0.29)	675	9.1 (0.39)	368	13.1 (0.78)	61,106	12.5 (0.12)
Non-Hispanic white.	5,727	68.1 (0.58)	2,210	66.9 (0.83)	766	[2]58.7 (1.36)	162,530	71.3 (0.18)
Non-Hispanic black.	2,219	20.3 (0.50)	927	21.3 (0.69)	418	[2]24.9 (1.17)	31,768	11.2 (0.13)
Non-Hispanic, all other races	228	2.7 (0.19)	87	2.7 (0.37)	46	[2]3.3 (0.50)	11,211	5.0 (0.08)
Poverty status (percent of poverty threshold)								
Less than 100%	3,998	38.4 (0.65)	1,406	[1]34.6 (0.95)	549	[2]31.3 (1.39)	34,361	10.6 (0.12)
100%–199%	3,036	30.9 (0.62)	1,142	[1]28.5 (0.95)	547	[2]33.9 (1.33)	47,881	15.7 (0.12)
200% or more.	2,779	30.7 (0.54)	1,351	[1]36.9 (0.93)	502	[2]34.8 (1.40)	184,373	73.7 (0.16)
Education								
No high school diploma	3,956	37.8 (0.56)	1,439	[1]34.0 (0.85)	714	[2]40.3 (1.34)	50,115	14.6 (0.11)
High school diploma or GED	3,202	33.7 (0.54)	1,349	[1]36.2 (0.87)	487	[2]33.1 (1.32)	77,734	29.5 (0.13)
More than a high school education.	2,497	26.9 (0.52)	1,050	28.3 (0.78)	331	[2]22.6 (1.18)	136,328	55.2 (0.16)
Missing	158	1.6 (0.15)	61	1.5 (0.20)	66	[2]3.9 (0.54)	2,438	0.8 (0.02)
Employment								
Employed in last week	908	9.9 (0.35)	590	[1]15.5 (0.64)	234	[2]15.3 (1.02)	216,780	82.2 (0.09)
Not employed in last week	8,902	90.1 (0.35)	3,299	[1]84.2 (0.65)	1,364	[2]84.7 (1.02)	49,465	17.7 (0.09)
Missing	3	*	10	*	0	*	370	0.1 (0.01)
Marital status								
Married.	3,521	36.6 (0.58)	1,726	[1]45.1 (0.93)	583	[2]38.1 (1.37)	159,995	60.3 (0.15)
Not married.	6,272	63.2 (0.58)	2,163	[1]54.7 (0.93)	1,007	[2]61.4 (1.37)	106,157	39.6 (0.15)
Missing	20	0.2 (0.05)	10	*	8	*	463	0.2 (0.01)
Health status								
Excellent or very good	1,182	12.5 (0.38)	593	[1]14.7 (0.63)	260	[2]16.8 (1.07)	180,510	69.5 (0.13)
Good.	2,465	25.9 (0.50)	1,016	26.4 (0.77)	421	[2]26.9 (1.23)	66,277	23.7 (0.11)
Fair or poor.	6,151	61.5 (0.56)	2,280	[1]58.6 (0.87)	909	[2]55.8 (1.44)	19,614	6.7 (0.06)
Missing	15	0.2 (0.05)	10	0.3 (0.08)	8	*	214	0.1 (0.01)

* Estimate does not meet NCHS standards of reliability.
[1]p value < 0.05 compared with concordant Yes–SSA and Yes–NHIS group.
[2]p value < 0.05 compared with concordant No–SSA and No–NHIS group.

NOTES: SSA is Social Security Administration. NHIS is National Health Interview Survey. SE is standard error.

SOURCE: NCHS, Linked National Health Interview Survey and Social Security Administration data.

Table 5. Weighted percentages and standard errors for selected characteristics for linkage-eligible National Health Interview Survey participants aged 18–64, by concordant–discordant groups (linked National Health Interview Survey and Social Security Administration Social Security Disability Insurance benefits data), 1998–2005

Characteristic	Yes–SSA and Yes–NHIS		Yes–SSA and No–NHIS		No–SSA and Yes–NHIS		No–SSA and No–NHIS	
	n	Percent (SE)	n	Percent (SE)	n	Percent (SE)	n	Percent (SE)
Total	5,471	100.0	3,777	100.0	1,549	100.0	270,871	100.0
Age group								
18–39	733	13.5 (0.52)	867	[1]22.7 (0.73)	398	[2]25.6 (1.21)	142,317	51.2 (0.13)
40–64	4,738	86.5 (0.52)	2,910	[1]77.3 (0.73)	1,151	[2]74.4 (1.21)	128,554	48.8 (0.13)
Sex								
Male	2,974	53.7 (0.71)	1,902	[1]50.7 (0.86)	786	[2]52.0 (1.36)	132,223	49.1 (0.08)
Female	2,497	46.3 (0.71)	1,875	[1]49.3 (0.86)	763	[2]48.0 (1.36)	138,648	50.9 (0.08)
Race and ethnicity								
Hispanic	700	6.7 (0.32)	578	[1]8.2 (0.41)	341	12.3 (0.81)	62,122	12.5 (0.12)
Non-Hispanic white.	3,657	75.1 (0.64)	2,310	[1]70.7 (0.80)	777	[2]60.9 (1.42)	164,346	71.1 (0.18)
Non-Hispanic black.	1,017	16.2 (0.55)	821	[1]19.1 (0.69)	385	[2]23.1 (1.19)	33,058	11.4 (0.13)
Non-Hispanic, all other races	97	2.0 (0.21)	68	2.0 (0.27)	46	[2]3.7 (0.58)	11,345	5.0 (0.08)
Poverty status (percent of poverty threshold)								
Less than 100%	1,592	27.2 (0.75)	1,178	[1]30.0 (0.92)	603	[2]35.5 (1.37)	36,891	11.3 (0.12)
100%–199%	1,865	33.7 (0.80)	1,163	[1]30.0 (0.94)	483	[2]31.4 (1.44)	49,048	15.9 (0.12)
200% or more.	2,013	39.1 (0.75)	1,436	40.0 (0.95)	464	[2]33.2 (1.40)	184,932	72.8 (0.16)
Education								
No high school diploma	1,770	30.2 (0.68)	1,235	30.7 (0.84)	686	[2]40.8 (1.37)	52,469	15.1 (0.11)
High school diploma or GED	1,873	35.0 (0.70)	1,357	37.0 (0.88)	468	[2]32.5 (1.32)	79,005	29.5 (0.13)
More than a high school education.	1,774	33.8 (0.70)	1,121	[1]30.7 (0.82)	349	[2]23.5 (1.20)	136,869	54.6 (0.16)
Missing	54	0.9 (0.13)	64	[1]1.6 (0.22)	46	[2]3.2 (0.51)	2,528	0.8 (0.02)
Employment								
Employed in last week	489	9.3 (0.44)	540	[1]15.2 (0.66)	197	[2]13.8 (1.03)	217,142	81.1 (0.10)
Not employed in last week	4,982	90.7 (0.44)	3,230	[1]84.6 (0.66)	1,352	[2]86.2 (1.03)	53,366	18.7 (0.09)
Missing	0	*	7	*	0	*	363	0.1 (0.01)
Marital status								
Married	2,544	47.1 (0.76)	1,713	45.8 (0.92)	501	[2]34.1 (1.39)	160,938	59.7 (0.15)
Not married.	2,919	52.8 (0.76)	2,055	54.0 (0.92)	1,039	[2]65.4 (1.40)	109,465	40.1 (0.15)
Missing	8	*	9	*	9	*	468	0.2 (0.01)
Health status								
Excellent or very good	536	10.3 (0.45)	529	[1]13.7 (0.62)	225	[2]15.1 (1.03)	181,132	68.7 (0.13)
Good.	1,329	24.9 (0.63)	1,018	[1]27.7 (0.81)	413	[2]27.1 (1.25)	67,335	23.7 (0.11)
Fair or poor.	3,601	64.7 (0.70)	2,220	[1]58.3 (0.86)	906	[2]57.6 (1.42)	22,182	7.5 (0.07)
Missing	5	*	10	*	5	*	222	0.1 (0.01)

* Estimate does not meet NCHS standards of reliability.
[1]p value < 0.05 compared with concordant Yes–SSA and Yes–NHIS group.
[2]p value < 0.05 compared with concordant No–SSA and No–NHIS group.

NOTES: SSA is Social Security Administration. NHIS is National Health Interview Survey. SE is standard error.

SOURCE: NCHS, Linked National Health Interview Survey and Social Security Administration data.

Table 6. Weighted percentages and standard errors for selected characteristics for linkage-eligible National Health Interview Survey participants aged 18–64, by concordant–discordant groups (linked National Health Interview Survey and Social Security Administration record of receipt of Supplemental Security Income benefits data), 1998–2005

Characteristic	Yes–SSA and Yes–NHIS		Yes–SSA and No–NHIS		No–SSA and Yes–NHIS		No–SSA and No–NHIS	
	n	Percent (SE)	n	Percent (SE)	n	Percent (SE)	n	Percent (SE)
Total	4,328	100.0	1,914	100.0	1,391	100.0	273,335	100.0
Age group								
18–39	1,576	37.4 (0.82)	740	40.0 (1.26)	421	[2]30.1 (1.40)	141,362	50.4 (0.13)
40–64	2,752	62.6 (0.82)	1,174	60.0 (1.26)	970	[2]69.9 (1.40)	131,973	49.6 (0.13)
Sex								
Male	1,660	39.6 (0.81)	816	[1]44.1 (1.27)	698	49.6 (1.47)	134,426	49.4 (0.08)
Female	2,668	60.4 (0.81)	1,098	[1]55.9 (1.27)	693	50.4 (1.47)	138,909	50.6 (0.08)
Race and ethnicity								
Hispanic	931	11.7 (0.49)	455	12.6 (0.64)	310	12.7 (0.85)	61,879	12.3 (0.12)
Non-Hispanic white.	2,086	58.8 (0.94)	838	[1]54.9 (1.23)	696	[2]60.6 (1.44)	167,159	71.5 (0.18)
Non-Hispanic black.	1,187	26.0 (0.82)	568	[1]29.0 (1.10)	342	[2]23.3 (1.23)	33,008	11.3 (0.13)
Non-Hispanic, all other races	124	3.5 (0.36)	53	3.4 (0.54)	43	[2]3.4 (0.50)	11,289	4.9 (0.08)
Poverty status (percent of poverty threshold)								
Less than 100%	2,560	58.0 (0.98)	978	[1]50.6 (1.41)	506	[2]33.6 (1.48)	35,744	10.8 (0.11)
100%–199%	1,152	26.6 (0.86)	554	28.1 (1.29)	460	[2]32.8 (1.45)	50,255	16.2 (0.12)
200% or more.	615	15.4 (0.68)	383	[1]21.3 (1.13)	425	[2]33.6 (1.49)	187,336	73.0 (0.16)
Education								
No high school diploma	2,233	49.3 (0.87)	901	[1]43.9 (1.27)	582	[2]38.9 (1.45)	52,189	14.9 (0.11)
High school diploma or GED	1,305	31.6 (0.80)	628	[1]34.7 (1.18)	450	[2]33.6 (1.40)	80,105	29.6 (0.13)
More than a high school education.	696	16.8 (0.66)	345	[1]19.3 (1.03)	307	[2]24.1 (1.29)	138,604	54.7 (0.16)
Missing	94	2.3 (0.29)	40	2.1 (0.35)	52	[2]3.4 (0.54)	2,437	0.8 (0.02)
Employment								
Employed in last week	404	10.1 (0.55)	315	[1]17.1 (0.97)	198	[2]15.3 (1.13)	217,236	80.3 (0.09)
Not employed in last week	3,922	89.9 (0.55)	1,594	[1]82.7 (0.97)	1,192	[2]84.6 (1.13)	55,747	19.6 (0.09)
Missing	2	*	5	*	1	*	352	0.1 (0.01)
Marital status								
Married	896	20.1 (0.75)	525	[1]26.3 (1.14)	448	[2]32.7 (1.44)	163,753	60.2 (0.15)
Not married.	3,422	79.6 (0.76)	1,381	[1]73.3 (1.14)	934	[2]66.7 (1.44)	109,116	39.7 (0.15)
Missing	10	*	8	*	9	*	466	0.2 (0.01)
Health status								
Excellent or very good	603	14.4 (0.59)	358	[1]18.7 (0.98)	235	[2]17.2 (1.11)	181,059	68.0 (0.13)
Good.	1,091	25.9 (0.77)	504	26.8 (1.14)	386	[2]29.3 (1.39)	67,945	23.7 (0.11)
Fair or poor.	2,625	59.4 (0.87)	1,050	[1]54.3 (1.28)	764	[2]53.0 (1.55)	24,114	8.2 (0.07)
Missing	9	*	2	*	6	*	217	0.1 (0.01)

* Estimate does not meet NCHS standards of reliability.
[1]p value < 0.05 compared with concordant Yes–SSA and Yes–NHIS group.
[2]p value < 0.05 compared with concordant No–SSA and No–NHIS group.

NOTES: SSA is Social Security Administration. NHIS is National Health Interview Survey. SE is standard error.

SOURCE: NCHS, Linked National Health Interview Survey and Social Security Administration data.

FIRST CLASS MAIL
POSTAGE & FEES PAID
CDC/NCHS
PERMIT NO. G-284

For more NCHS NHSRs, visit:
https://www.cdc.gov/nchs/products/nhsr.htm.

National Health Statistics Reports ■ Number 131 ■ November 4, 2019

Suggested citation

Copyright information

National Center for Health Statistics

Jennifer H. Madans, Ph.D., *Acting Director*
Amy M. Branum, Ph.D., *Acting Associate Director for Science*

Division of Analysis and Epidemiology

Irma E. Arispe, Ph.D., *Director*
Kevin C. Heslin, Ph.D., *Associate Director for Science*