Published in final edited form as:

Epidemiology. 2018 September ; 29(5): 716–720. doi:10.1097/EDE.0000000000000861.

Misclassification of Rental Assistance in the National Health Interview Survey:

Evidence and Implications

Michel Boudreaux^a, Andrew Fenelon^a, Natalie Slopen^b

^aDepartment of Health Services Administration, School of Public Health, University of Maryland, MD

^bDepartment of Epidemiology and Biostatistics, School of Public Health, University of Maryland, MD.

Abstract

Background: Federal surveys could play a role in measuring the association of rental assistance and health and in identifying the health needs of the assisted population. However, self-reports of rental assistance could be biased. Our objective was to assess the accuracy of reported rental assistance in the National Health Interview Survey (NHIS).

Methods: We conducted a record-check study of reports of US Department of Housing and Urban Development rental assistance in the 2004–2012 NHIS, using survey responses linked to administrative records. Misclassification measures were limited to the false-negative rate because the survey ascertained participation in all rental assistance programs, but the administrative data pertained only to US Department of Housing and Urban Development. False-negative rates were calculated for the total population, for sociodemographic subgroups, across levels of self-reported health status, and for specific assistance types (Housing Choice Vouchers, Public Housing, and Multifamily Housing).

Results: We estimated a false-negative rate of 22.6%. Misclassification was higher among Public Housing residents compared to those receiving other forms of assistance, even after controlling for sociodemographics. Rates varied across region and other demographics. Those self-reporting fair or poor health were less likely to misreport assistance compared with those in better health, but the difference was explained by covariates. Misreporting assistance had little independent impact on the adjusted association of assistance and health.

Conclusions: False-negative reporting of rental assistance is moderately high in the NHIS, but we did not find evidence that it independently biased estimates of the association of health and rental assistance.

Correspondence: Michel Boudreaux, Department of Health Services Administration School of Public Health, University of Maryland, 4200 Valley Drive, No. 3310A College Park, MD 20742. mhb@umd.edu.

Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

The authors report no conflicts of interest.

Keywords

Misclassification; National Health Interview Survey; Rental assistance

Rental assistance from the US Department of Housing and Urban Development (HUD) has the potential to improve the health of low-income populations by improving housing affordability, quality, and stability.^{1–3} HUD rental assistance takes three primary forms: Housing Choice Vouchers (rent subsidies paid directly to landlords on behalf of tenants renting in the private market), Public Housing (developments available only to assisted families), and Multifamily Housing (developments where some units are reserved for assisted families).⁴

Federal surveys could play an important role in estimating the effect of HUD assistance on health and in identifying the health needs of the assisted population.^{5,6} Unlike administrative records, surveys provide a rich set of health outcomes and other important covariates, sample nonassisted populations that can be used as comparison groups, and can be easily obtained by the average researcher.

However, self-reports of rental assistance are considered to be inaccurate. An analysis of the 1989 American Housing Survey suggested that 11% of truly assisted households did not report assistance and 19% of nonassisted household reported assistance.⁷ More recent evidence from the Current Population Survey suggests that 36% of known HUD participants, identified from linked administrative data, do not report assistance.⁸ Many have argued that the misclassification of rental assistance poses a serious threat to the validity of estimates of the association of rental assistance and health, but the evidence on reporting errors of HUD assistance is limited to just a few surveys, and no studies estimate reporting errors in a federally conducted health survey.^{2,9–11} A complete review of the literature on public program reporting is presented in eAppendix, Section 1; <http://links.lww.com/EDE/B358>.

The goal of this study was to characterize the accuracy of self-reported HUD assistance in the National Health Interview Survey (NHIS) using individually linked administrative records. We make three contributions: (1) we present the first estimates of misclassification of HUD participation in a widely used health survey; (2) we examine misclassification across HUD's three major programs; and (3) we measure the extent to which misclassification leads to bias in estimates of the association of assistance and health.

METHODS

Data

We conducted a record-check study using a version of the 2004–2012 NHIS that was linked to HUD administrative records.¹² We accessed these data through an agreement with the National Center for Health Statistics (NCHS) and conducted analyses at a Federal Statistical Research Data Center. The University of Maryland Institutional Review Board reviewed and approved the study. The administrative elements we used were enrollment on the date of

interview, duration of the current participation episode, and program type (Housing Choice Vouchers, Multifamily Housing, or Public Housing).

The NHIS is an in-person household survey. The complex sample is representative of the civilian noninstitutionalized US population.¹³ The NHIS achieves approximately an 80% household response rate. The total sample size in 2004–2012 was $n = 807,253$ (see eAppendix Figure 1; <http://links.lww.com/EDE/B358> for the sample flow diagram). We used a harmonized version of NHIS variables to ease coding over time.¹⁴

The question on rental assistance is asked only of respondents who indicate their dwelling was rented. The question asks, “[Are you/Is anyone in your family] paying lower rent because the Federal, State, or local government is paying part of the cost?”.¹⁵ Respondents may answer “yes,” “no,” “do not know,” or “refuse.”

To validate the question, we compared responses to individually linked HUD records that indicated participation on the date of interview.¹² Consistent with other studies, we treated the administrative record as the gold standard given that it serves official purposes. However, administrative records are not free from error.¹²

NHIS respondents who provided complete information on the variables used in the linking process and who did not refuse the rental assistance question were eligible to be linked. Those eligible for linkage were deterministically matched using 9- or 4-digit social security numbers (SSN; depending on the year), month and year of birth, and sex.¹² However, the NHIS did not attempt to collect SSNs from every subject, and the rules determining which subjects were asked for SSNs changed over time.¹² To account for differential availability of SSNs across years, we used only the “sample adult” and “sample child” files (everyone in these files was asked for SSNs). The files include detailed information on one randomly selected adult and one randomly selected child per household. There were 356,204 sample cases in 2004–2012.

After excluding the 177 cases that refused the rental assistance question, 42% to 62% of remaining sample adults and 25% to 36% of remaining sample children were linkage eligible (164,458 cases). Imperfect linkage eligibility was primarily caused by incomplete SSNs or birthdates. As shown in eAppendix Table 1; <http://links.lww.com/EDE/B358>, the probability of being linkage eligible varied across sociodemographics and self-reported rental assistance status. Importantly, 54.9% of the 19,052 self-reported assistance cases were linkage eligible compared to 45.4% of 110,457 nonassisted renters and 45.9% of 226,156 nonrenters. These differences remained after adjusting for sociodemographics, suggesting that validation estimates from the linked sample would not be representative of the full NHIS.

Weighting

NCHS developed adjusted weights for the linkage-eligible sample to ensure that it would be representative of the target population.¹² However, estimates in eAppendix Table 3; <http://links.lww.com/EDE/B358> suggest that the adjusted weight does not produce fully representative estimates. In the full sample using the original weight, 3.6% of the population

is estimated to report assistance compared to 4.6% of linkage-eligible cases weighted using the adjusted weight.

To account for the nonrepresentativeness of the linkage-eligible sample, we created new weights. The new weights were designed to ensure that the linkage-eligible sample was representative of the NHIS target population on every variable included in our analysis. This new weight (called the modeled weight below) was created by predicting linkage eligibility using the full set of variables included in our analysis. We then adjusted the original weight of linkage-eligible cases by the inverse of the predicted probability of linkage eligibility. As shown in eAppendix Table 3; <http://links.lww.com/EDE/B358>, the modeled weight perfectly reproduces the original distributions.

False-negative Reports

We used the linked data and the modeled weight to calculate false-negative rates for the total population and for subgroups. The false-negative rate was defined as the percentage of people who reported they did not receive assistance, among all people that the administrative data indicated were enrolled on the date of interview. The denominator of our preferred measure included all renters and nonrenters. We also report how the false-negative rate changes when we exclude nonrenters and those unaware of their assistance status and when we use the NCHS-adjusted weights.

Because the survey measures participation in any rental assistance program, but the validation data pertained only to HUD, we could not calculate other classification measures such as sensitivity. In eAppendix Table 4; <http://links.lww.com/EDE/B358>, we show that on an aggregate basis, the NHIS estimates more people with assistance than is suggested by HUD administrative counts. However, because we did not observe true assistance status for all programs, we cannot determine what share of the overcount stems from misreporting.

Subgroups included program type and duration of the current assistance episode, demographics (region, age, race, and sex), poverty level, and self-reported health. More subgroups are considered in eAppendix Table 5; <http://links.lww.com/EDE/B358>. We report both unadjusted false-negative rates and adjusted percentage point differences across subgroups (i.e., rate differences derived from a logistic regression).

Reporting Error and Self-reported Health Status

We also examined the extent to which false-negative reports biased estimates of the association of HUD assistance and self-reported health (“Would you say your health in general is excellent, very good, good, fair, or poor?”), dichotomized as fair/poor versus better. Self-reported health has been shown to be a valid and reliable measure of physical and emotional health.¹⁶ However, our results might not generalize to specific health conditions. We began by estimating the association of fair/poor health with self-reported rental assistance, adjusting for observed controls. We then repeated the analysis after substituting self-reported assistance with the administrative indicator and finally with self-reported rental assistance edited to “yes” if the linked HUD record indicated participation. The edited variable only changed known false-negative responses to true positives but did not alter potential false-positive responses.

The linkage-eligible sample included 164,458 sample adult and sample children of which 8,052 were indicated to be HUD participants on the day of interview. Standard errors accounted for the complex sample design.

RESULTS

We report false-negative rates by denominator and weight in Table 1. The unweighted false-negative rate was 13.8% among renters who provided a yes/no response to the assistance question. Including nonrenters and do not know assistance responses increased the unweighted rate to 17.5%. The NCHS-adjusted weight estimated an 18.9% false-negative rate in the population that included nonrenters and do not know assistance responses. Using the modeled weight, we estimated a false-negative rate of 22.6% when including nonrenters and do not know responses.

In Table 2 we report false-negative rates by subgroup using the modeled weight. (A more complete set of subgroups is presented in eAppendix Table 5; <http://links.lww.com/EDE/B358> along with tests of association.) There was substantial variation across HUD's three programs. The false-negative rate was 20.6% for Housing Choice Vouchers, 17.9% for Multifamily Housing, and 32.9% for Public Housing. Public Housing residents remained more likely than the voucher population to be misclassified after controlling for covariates. There was no meaningful variation by duration of the current episode.

Residents in the West had an unadjusted false-negative rate of 14.4%, and residents in the Northeast had a false-negative rate of 27.4%. There was a steep gradient in false-negative reporting by poverty: after adjustment, those above 150% of the poverty line were 23.4% points more likely not to report assistance compared to those with family incomes below 50% of poverty.

There was a 5-point difference in the false-negative rate among those with fair/poor health compared with those with good health or better. However, this difference was largely explained by covariates (Table 2), suggesting that reporting error of rental assistance does not independently bias the association of assistance and self-reported health. After adjusting for covariates, the difference in self-reported health between assisted and nonassisted respondents was nearly equal for the self-reported, administrative, and edited measures (eAppendix Table 9; <http://links.lww.com/EDE/B358>). Using the self-reported measure, there was a 2.1 (95% confidence interval [CI] = 1.7, 2.6) percentage-point difference in self-reported health by assistance status compared to a 2.0 (95% CI = 1.3, 2.7) percentage-point difference using the HUD indicator and a 2.1 (95% CI = 1.5, 2.7) percentage-point difference using the edited variable.

DISCUSSION

Our preferred estimates suggest a false-negative rate for HUD assistance of 22.6% in the 2004–2012 NHIS. In absolute terms, this suggests a fair amount of misreporting. However, the false-negative rate of HUD assistance in the NHIS is lower than in the Current Population Survey and lower than for other public programs in the NHIS.^{8,17,18}

Misreporting of HUD assistance varied across subgroups. After adjustment, the false-negative rate for Public Housing residents was 11.6 percentage-points higher than it was for the voucher population. A previous study of HUD participants found that participants in all programs are well-informed about their benefits. However, many understand the role the government plays in subsidizing their rent differently than survey questions intend.¹⁰ For a Public Housing resident who understands that the government owns their unit, a true-positive response requires recognizing that the rent is set below cost and that the loss the government incurs is equivalent to the “pays part of the cost” language used in the question. Interaction with a private landlord through Housing Choice Vouchers might provide a better signal that the government is paying.

Like other public program questions, people in higher-income families were more likely to be misclassified than lower-income people.^{8,18–19} This could signal that respondents with more contact with the safety net are more knowledgeable about their benefits or are less affected by social desirability.

Although the level of misclassification is concerning, misreporting did not appear to substantially bias estimates of the association between assistance and health status. After controlling for covariates, the association of self-reported health and assistance was similar using the self-reported indicator compared to the administrative indicator. This suggests that data users can have some confidence in estimating correlations between assistance and health using multivariable regression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We are grateful to Dieudonne Nahigombeye of the National Center for Health Statistics and Sandy Dietrich of the Census Bureau for data assistance.

Replication data: This study is based on restricted use data from the National Center for Health Statistics (NCHS). NCHS does not permit public dissemination of the data, but they are obtainable through a restricted data use agreement.

Work for this project was supported by a seed grant from the Maryland Population Research Center at the University of Maryland, College Park.

REFERENCES

1. Helms V, Sperling J, Steffen B. A health picture of HUD-assisted adults, 2006–2012. U.S. Department of Housing and Urban Development, Office of Policy Development and Research Available at: <https://www.huduser.gov/portal/sites/default/files/pdf/Health-Picture-of-HUD.pdf> Accessed December 15, 2017.
2. Kushel MB, Gupta R, Gee L, Haas JS. Housing instability and food insecurity as barriers to health care among low-income Americans. *J Gen Intern Med*. 2006;21:71–77. [PubMed: 16423128]
3. Newman S, Holupka C. Housing affordability and child well-being. *Housing Policy Debate*. 2015;25:116–151.
4. Center for Budget and Policy Priorities. Fact sheet: Federal rental assistance. Available at: https://www.cbpp.org/sites/default/files/atoms/files/4-13-11_hous-US.pdf. Accessed December 15, 2017.

5. Fenelon A, Mayne P, Simon A, et al. Housing assistance programs and adult health in the United States." Am J Public Health. 2017;107:571–578. [PubMed: 28207335]
6. Simon AE, Fenelon A, Helms V Lloyd PC, Rossen LM. HUD housing assistance associated with lower uninsurance rates and unmet medical need. Health Aff (Millwood). 2017;36:1016–1023. [PubMed: 28583959]
7. Shroder M Does housing assistance perversely affect self-sufficiency? A review essay. J Housing Econ. 2002;11:381–417.
8. Meyer B, Mittag N. Using linked survey and administrative data to better measure income: implications for poverty, program effectiveness, and holes in the safety net. National Bureau of Economic Research. 2015; Working Paper No. 21676.
9. Helms VE, King BA, Ashley PJ. Cigarette smoking and adverse health outcomes among adults receiving federal housing assistance. Prev Med. 2017;99:171–177. [PubMed: 28192095]
10. Gordon E, Chipunga S, Bagley L, Zanakos S. Improving housing subsidy surveys: data collection techniques for identifying the housing subsidy status of survey respondents. U.S. Department of Housing and Urban Development, Office of Policy Development and Research Available at: <https://www.huduser.gov/publications/pdf/datacollect.pdf>. Accessed December 15, 2017
11. Susin S Longitudinal outcomes of subsidized housing recipients in matched survey and administrative data. Cityscape. 2005;189–218.
12. Lloyd P, Helms V. NCHS-HUD linked data: analytic considerations and guidelines. National Center for Health Statistics. Available at: https://www.cdc.gov/nchs/data/datalinkage/nchs_hud_analytic_considerations.pdf. Accessed December 15, 2015.
13. Parsons V Moriarity C, Jonas K, More T, Davis K, Tompkins L. Design and estimation for the National Health Interview Survey, 2006–2015. National Center for Health Statistics. Vital Health Stat 2. 2014;(165):1–53.
14. Blewett L, Drew J, Griffin R, King M, Williams K. IPUMS Health Surveys: National Health Interview Survey, Version 6.2. Minneapolis: University of Minnesota; 2016.
15. National Center for Health Statistics. 2010 National Health Interview Survey: Family. 2011 Available at: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2010/familyxx_layout.pdf Accessed December 15, 2017.
16. DeSalvo KB, Fisher WP, Tran K, Bloser N, Merrill W, Peabody J. Assessing measurement properties of two single-item general health measures. Qual Life Res. 2006;15:191–201. [PubMed: 16468076]
17. Call KT, Davern ME, Klerman JA, Lynch V Comparing errors in Medicaid reporting across surveys: evidence to date. Health Serv Res. 2013;48(2 Pt 1):652–664. [PubMed: 22816493]
18. University of Minnesota's State Health Access Data Assistance Center, the Centers for Medicare and Medicaid Services, the Department of Health and Human Services Assistant Secretary for Planning and Evaluation, The National Center for Health Statistics, and the U.S Census Bureau (SNACC). Phase IV research results. Available at: <http://www.census.gov/did/www/snacc/>. Accessed December 15, 2017.
19. Boudreaux MH, Call KT, Turner J, Fried B, O'Hara B. Measurement error in public health insurance reporting in the American Community Survey: evidence from record linkage. Health Serv Res. 2015;50:1973–1995. [PubMed: 25865628]

TABLE 1.

False-negative Rates for Rental Assistance by Denominator and Weight, 2004–2012 NHIS-HUD

	False-negative Rate (95% Confidence Interval)
Unweighted	
Do not know rental assistance and non-renters excluded	13.8 (13.1, 14.6)
Assign do not know rental assistance to no assistance	14.4 (13.6, 15.2)
Assign do not know rental assistance and non-renters to no assistance	17.5 (16.7, 18.3)
NCHS-adjusted weighted	
Do not know assistance and non-renters excluded	15.1 (13.3, 16.9)
Assign do not know rental assistance to no assistance	15.4 (13.6, 17.2)
Assign do not know rental assistance and non-renters as no assistance	18.9 (17.2, 20.6)
Modeled weight	
Do not know assistance and non-renters excluded	18.0 (15.9, 20.1)
Assign do not know rental assistance to no	18.3 (16.2, 20.5)
Assign do not know rental assistance and non-renters as no assistance	22.6 (20.6, 24.7)

2004–2012 National Health Interview Survey–Housing and Urban Development (NHIS-HUD) Linked Files, Sample Adults and Sample Children that are enrolled in HUD on the date of interview. There are 7,711 observations using the most restrictive denominator, 7,738 observation when do not know rental assistance is considered a false-negative report, and 8,052 observations when non-renters are added as false-negative negative report, and 8,052 observations when non-renters are reports. Confidence intervals based on linearized standard errors that account for the sample design.

NCHS indicates National Center for Health Statistics.

TABLE 2.

False-negative Rates of Self-Reported Rental Assistance and Rate Differences by Subgroup, NHIS-HUD 2004–2012

	Unadjusted	Adjusted
	False-negative Rate (95% Confidence Interval)	Rate Difference (95% Confidence Interval)
Overall	22.6 (20.6, 24.7)	
Assistance type		
Housing choice vouchers	20.6 (18.0, 23.1)	(REF)
Multi family housing	18.0 (15.0, 20.9)	−0.5 (−4.0, 3.1)
Public housing	32.9 (27.1, 38.7)	11.6 (6.2, 16.9)
Duration of current HUD episode		
Less than 1 year	23.2 (19.6, 26.7)	(REF)
1–3 years	22.7 (19.4, 26.1)	−1.1 (−4.7, 2.6)
3 or more years	22.3 (19.8, 24.8)	−1.4 (−4.8, 2.0)
Region		
Northeast	27.4 (22.4, 32.5)	(REF)
Midwest	19.1 (16.3, 22.0)	−2.5 (−8.1, 3.1)
South	25.8 (22.7, 28.9)	0.7 (−4.4, 5.9)
West	14.4 (9.9, 18.9)	−8.8 (−14.4, −3.2)
Sex		
Female	22.7 (20.4, 24.9)	(REF)
Male	22.6 (19.6, 25.5)	−1.9 (−4.5, 0.7)
Age		
0–5	22.1 (16.8, 27.4)	(REF)
6–12	22.5 (17.6, 27.4)	0.6 (−5.3, 6.6)
13–17	27.3 (22.3, 32.3)	3.1 (−3.6, 9.7)
18–24	29.2 (24.2, 34.2)	4.4 (−0.9, 9.7)
25–44	22.0 (19.1, 24.8)	−0.3 (−4.3, 3.7)
45–64	20.2 (17.4, 23.0)	−0.9 (−6.1, 4.4)
65+	19.4 (16.0, 22.9)	−0.1 (−6.6, 6.5)
Race		
Non-Hispanic White	18.5 (15.3, 21.7)	(REF)
Non-Hispanic Black	25.2 (22.5, 27.9)	3.6 (0.2, 7.0)
Non-Hispanic other/multiple	15.8 (7.9, 23.6)	−2.5 (−8.7, 3.7)
Hispanic	25.2 (20.8, 29.6)	3.6 (−1.4, 8.6)
Poverty status (% FPL)		
0–49	17.5 (14.7, 20.4)	(REF)
50–99	17.3 (14.8, 19.9)	1.0 (−2.0, 4.0)
100–149	24.7 (20.3, 29.0)	7.4 (3.0, 11.8)
150+	45.4 (40.4, 50.5)	23.4 (18.0, 28.9)

	Unadjusted	Adjusted
	False-negative Rate (95% Confidence Interval)	Rate Difference (95% Confidence Interval)
Health status		
Good or better	23.9 (21.5, 26.2)	(REF)
Fair or poor	18.5 (16.1, 20.8)	−1.5 (−4.1, 1.2)

2004–2012 National Health Interview Survey-Housing and Urban Development (NHIS-HUD) Linked Files, Sample Children and Sample Adults. Estimates based on cases enrolled in HUD at the time of interview, including non-renters and do not know rental assistance responses. All estimates are weighted using the modeled weight (see text) and reported in percentage points. Rate differences were obtained from a logistic regression that included all covariates listed in the table in addition to survey year, family structure, family size, family-level employment, family-level education, and any medical provider visit (complete results are in the eAppendix). Confidence intervals are based on linearized standard errors.

FPL indicates federal poverty level; REF, referent groups.