

© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12528
RESEARCH BRIEF

Redaction of Substance Abuse Claims in Medicare Research Files Affects Spending Outcomes for Nearly One in Five Beneficiaries with Serious Mental Illness

Pamela Roberto, Nicole Brandt, Eberechukwu Onukwugha, and Bruce Stuart

Objective. To assess the impact of substance abuse claims redaction on Medicare spending estimates for beneficiaries with serious mental illness.

Data Sources. The 2012 claims and unredacted beneficiary-level Medicare spending totals from CMS's Chronic Conditions Warehouse.

Study Design. We identified beneficiaries with claims affected by the redaction by comparing claims-based spending estimates to unredacted spending totals. Differences in characteristics of beneficiaries with and without redacted claims were examined in bivariate analyses.

Principal Findings. Claims-based spending totals differed from unredacted totals for 19.7 percent of the cohort. Part A spending for those with redacted claims was underreported by 57.0 percent. Characteristics of beneficiaries with and without redacted claims differed significantly.

Conclusions. Researchers who rely on Medicare claims to analyze spending outcomes for beneficiaries with serious mental illness should be aware of the potential for bias due to nonrandom redaction of substance abuse data.

Key Words. Substance abuse claims redaction, Medicare spending, selection bias

At the behest of the Substance Abuse and Mental Health Services Administration (SAMHSA), the Centers for Medicare and Medicaid Services (CMS) currently redacts all claims containing substance abuse diagnosis or treatment information from the Medicare and Medicaid research identifiable files. According to data published by a CMS contractor, the impact of the claims redaction is limited in scope—affecting between 3 and 4

percent of the total Medicaid and 2–3 percent of the total Medicare population—although the share of beneficiaries with redacted claims has grown larger every year (Barosso 2015). On an aggregate level, the share of impacted claims appears minor, affecting less than 2 percent of total Medicare and Medicaid claims processed annually. However, certain types of services—particularly inpatient care—are disproportionately affected, with 7–8 percent of total inpatient claims missing for both the entire Medicare and Medicaid populations.

Although the redaction of substance abuse data affects only a small share of total Medicare and Medicaid claims, the practice is concerning to health services researchers for at least three reasons. First, the change in policy was poorly communicated. CMS began redacting all claims data released to researchers in late 2013 but did not publicly notify the research community until more than a year later (Frakt 2015). Second, it is not just the information pertaining to substance abuse on a claim that is redacted, but rather the entire claim. As most redacted claims are deleted because of a secondary diagnosis of substance abuse disorder (Frakt and Bagley 2015), important information regarding primary diagnoses and procedures may also be suppressed. Third, the redacted claims are not missing at random, which increases the potential for bias in studies that rely on claims data to generate utilization or spending measures, select cohort members based on diagnostic or procedural codes, or adjust for differences in comorbidity burden. The likelihood of bias resulting from this nonrandom claims deletion increases when studying populations with high rates of comorbid substance abuse, including those with costly chronic conditions such as HIV, hepatitis C, and depression (Rough et al. 2016). To better understand the impact of the substance abuse data redaction on claim-based estimates of Medicare spending, we assessed the accuracy of Parts A and B spending totals derived from the claims for a cohort of beneficiaries with serious mental illness, a population with an above-average rate of comorbid substance abuse disorders (Hartz et al. 2014).

Address correspondence to Pamela Roberto, M.P.P., Department of Pharmaceutical Health Services Research, University of Maryland School of Pharmacy, 220 Arch Street, 12th Floor, Baltimore, MD 21201; e-mails: pamroberto@umaryland.edu; proberto@phrma.org. Pamela Roberto, M.P.P., is also with the Policy & Research Department, PhRMA, Washington, DC. Nicole Brandt, Pharm.D., is with the Department of Pharmacy Practice and Science, University of Maryland School of Pharmacy, Baltimore, MD. Eberchukwu Onukwugha, Ph.D., M.S. and Bruce Stuart, Ph.D., are with the Department of Pharmaceutical Health Services Research, University of Maryland School of Pharmacy, Baltimore, MD.

METHODS

Data and Sample

Enrollment, demographic, and claims data were obtained for a 5 percent random sample of Medicare beneficiaries from the CMS Chronic Conditions Data Warehouse (CCW) for 2012, including summarized spending totals for each beneficiary from the Cost and Use file. Spending totals in the Cost and Use file are calculated by a CMS contractor using the full set of Medicare claims and are therefore unaffected by the claims redaction policy (ResDAC 2015, personal communication). The CCW data also contain chronic condition flags, generated from claims-based algorithms, which identify beneficiaries diagnosed with various chronic physical and mental illnesses.

Our cohort was limited to individuals who were continuously enrolled in Medicare Parts A, B, and D throughout 2012. Beneficiaries with CCW chronic condition flags for bipolar disorder, personality disorder, or schizophrenia or other psychoses (Buccaneer Computer Systems and Services, Inc 2016) and at least two antipsychotic prescription fills were identified as having serious mental illness ($N = 38,005$). Nursing home residents, Medicare Advantage enrollees, and beneficiaries without Part D prescription drug coverage were excluded from the analysis.

Study Design

Using algorithms specified in the CCW Medicare Administrative Data User Guide (Buccaneer Computer Systems and Services, Inc 2015), spending totals derived from the claims for all Parts A and B service types were compared to the unredacted summary spending totals reported in the Cost and Use file for each beneficiary. Part A spending included acute and other inpatient hospital, skilled nursing facility, hospice, and home health care, while Part B spending was comprised of hospital outpatient, ambulatory surgical center, anesthesia, Part B drug, physician office, evaluation and management, dialysis, imaging, laboratory, durable medical equipment, other procedures, and other Part B carrier spending. Beneficiaries whose claims-based spending totals for each service type did not match their corresponding Cost and Use summary spending totals for that service type were classified as having claims affected by the redaction. We also compared Part D prescription drug spending totals from the claims to the Cost and Use summary totals as a falsification test. Because Part D claims are not affected by the redaction policy, claims-based spending totals should not

differ from the Cost and Use summary totals for prescription drugs covered under Part D. Differences in the characteristics of beneficiaries with and without redacted claims were assessed using t-tests and chi-squared tests.

RESULTS

Claims-based spending totals differed from Cost and Use summary totals for 19.7 percent of the cohort (Table 1). Nearly twice as many beneficiaries were affected by redaction of Part B claims than Part A claims (17.4 percent vs. 9.5 percent). Among beneficiaries with redacted claims for Part A-covered services, the majority involved acute and/or other inpatient hospital care. For those with redacted Part B claims, evaluation and management and hospital outpatient care were the most commonly affected services. Less than 1 percent of beneficiaries had redacted Part B claims for imaging, other procedures,

Table 1: Number and Share of Beneficiaries with Serious Mental Illness (N = 38,005) with Redacted Medicare Part A and Part B Claims, by Type of Service

	<i>Number with Redacted Claims</i>	<i>% with Redacted Claims</i>
Any Part A or Part B claims	7,493	19.7
Any Part A claims	3,608	9.5
Acute inpatient	2,112	5.6
Other inpatient	2,078	5.5
Home health	126	0.3
Skilled nursing facility	63	0.2
Hospice	0	0.0
Any Part B claims	6,613	17.4
Evaluation and management	4,390	11.6
Hospital outpatient	3,343	8.8
Laboratory or test	1,823	4.8
Physician office	1,326	3.5
Other Part B carrier services	685	1.8
Other procedures	251	0.7
Imaging	139	0.4
Part B drugs	143	0.4
Anesthesia	39	0.1
Durable medical equipment	*	*
Ambulatory surgical center	*	*
Dialysis	*	*
Any Part D claims	0	0.0

*Cell sizes with counts less than 11 were suppressed in accordance with the Centers for Medicare and Medicaid guidelines.

prescription drugs, anesthesia, durable medical equipment, ambulatory surgical center, or dialysis care. Part D prescription drug claims for all beneficiaries were unaffected by the redaction.

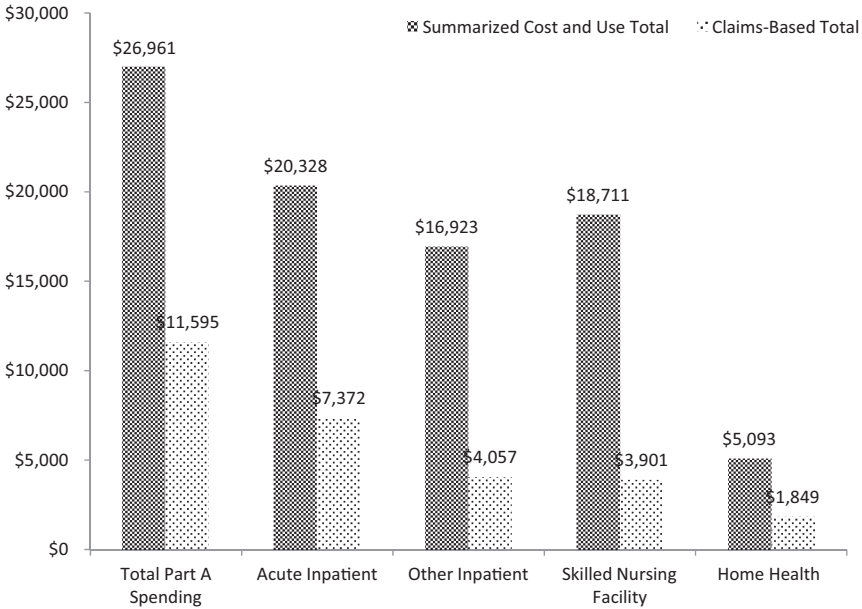
Based on the summarized spending totals reported in the Cost and Use file, average spending among beneficiaries with redacted claims far exceeded that of beneficiaries without redacted claims for every Part A and Part B service type (Table 2). For example, average acute inpatient expenditures incurred by beneficiaries with redacted claims were nearly 7 times greater than those of beneficiaries whose claims were unaffected (\$20,328 vs. \$2,905). Among beneficiaries with redacted claims, the impact of missing data on the claims-based spending measures, while nontrivial for Part B spending, was of considerable magnitude for Part A spending (Figures 1 and 2). On average, Part A spending for affected beneficiaries totaled \$26,961 in the summarized Cost and Use file, only \$11,595, or 43.0 percent, of which was observable in

Table 2: Mean Total Medicare Spending for Beneficiaries with Serious Mental Illness ($N = 38,005$) Reported in the Summarized Cost and Use File, by Claims Redaction Status

	<i>Beneficiaries with Claims Affected by the Redaction</i>	<i>Beneficiaries with No Redacted Claims</i>
Total Part A spending	\$26,960.78	\$6,501.57
Acute inpatient	\$20,328.09	\$2,904.88
Other inpatient	\$16,923.21	\$1,797.92
Skilled nursing facility	\$18,710.96	\$941.89
Home health	\$5,092.92	\$817.01
Hospice	N/A	\$143.29
Total Part B spending	\$8,111.24	\$4,786.12
Hospital outpatient	\$3,958.28	\$1,826.62
Evaluation and management	\$2,232.15	\$941.78
Other Part B carrier services	\$1,841.09	\$333.18
Laboratory or test	\$1,010.26	\$287.79
Physician office	\$830.11	\$438.48
Other procedures	\$806.71	\$389.99
Imaging	\$464.03	\$174.12
Anesthesia	\$460.81	\$55.31
Part B drugs	\$299.12	\$161.54
Ambulatory surgical center	*	*
Dialysis	*	*
Durable medical equipment	*	*

Notes. There were no beneficiaries whose claims for hospice care were affected by the redaction.
*Cell sizes with counts less than 11 were suppressed in accordance with the Centers for Medicare and Medicaid guidelines.

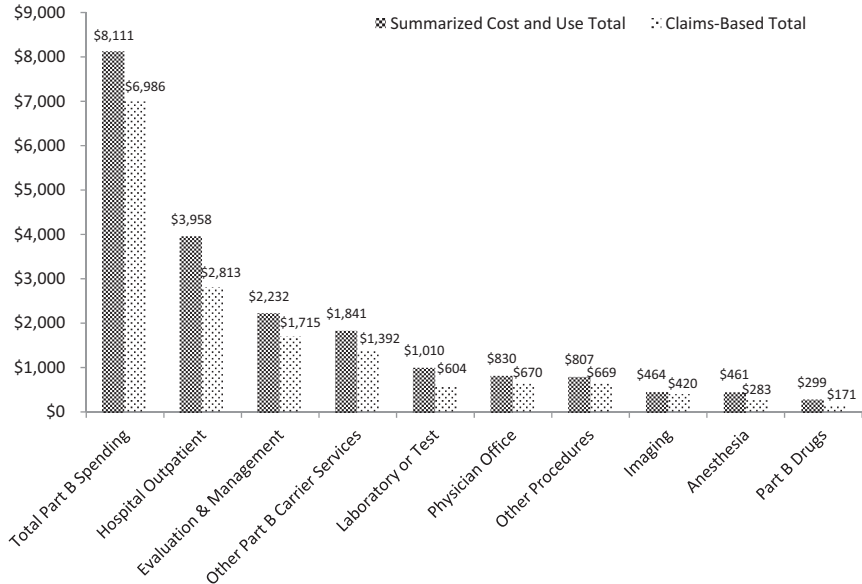
Figure 1: Mean Claims-Based Spending Totals versus Summarized Cost and Use Spending Totals for Medicare Part A Services for Beneficiaries with Serious Mental Illness with Redacted Claims



the claims. This pattern held for all Part A-covered service types, with the share of total spending present in the claims ranging from a low of 20.8 percent for skilled nursing facility care to a high of 36.3 percent for home health care. Nearly half (49.6 percent) of beneficiaries whose Part A claims were affected by the redaction had Part A spending reported in the summary Cost and Use file, but no corresponding spending in the Part A claims data, indicating that all of their Part A claims had been suppressed. Spending totals for Part B claims were less drastically affected: among beneficiaries with redacted data, 86.1 percent of total Part B spending remained present in the claims.

Significant differences were observed in the characteristics of beneficiaries with and without redacted claims (Table 3). On average, beneficiaries with redacted claims were younger (47.7 vs. 54.3 years of age), more likely to be male (53.5 percent vs. 47.7 percent), more likely to receive the Part D low-income subsidy (92.4 percent vs. 86.2 percent), more likely to live in the Northeast and Western regions of the United States, and less likely to live in the South. While there were no differences in the share of whites and Hispanics with

Figure 2: Mean Claims-Based Spending Totals versus Summarized Cost and Use Spending Totals for Medicare Part B Services for Beneficiaries with Serious Mental Illness with Redacted Claims



redacted claims, blacks were more likely, and beneficiaries of other races less likely, to have claims impacted by the redaction. Schizophrenia and other psychoses were less prevalent among beneficiaries affected by the redaction (62.9 percent vs. 66.9 percent), but the prevalence of bipolar disorder and personality disorder was substantially greater (71.4 percent vs. 50.4 percent and 23.2 percent vs. 7.6 percent, respectively). Beneficiaries with redacted claims also had significantly higher comorbidity burden, as measured by the Charlson comorbidity index, and significantly higher Medicare Parts A, B, and D spending.

DISCUSSION

Health service researchers often construct beneficiary-level utilization and spending measures directly from medical claims data and commonly rely on claim-based information, such as diagnostic and procedural codes, to identify disease cohorts and adjust for comorbidity burden. Incomplete data can introduce bias into such studies if the data are systematically missing for reasons related to the

Table 3: Characteristics of Beneficiaries with Serious Mental Illness with and without Redacted Claims

	Full Sample (N = 38,005)	With Redacted Claims (N = 7,493)	Without Redacted Claims (N = 30,512)
Age, mean (SD)	53.0 (15.2)	47.7 (12.6)	54.3 (15.6)***
Male (%)	47.4	53.5	45.9***
Race/ethnicity (%)			
White	75.2	74.6	75.3
Black	17.4	18.4	17.2*
Hispanic	3.7	3.9	3.6
Other	3.8	3.1	3.9**
Low-income subsidy recipient (%)	87.5	92.4	86.2***
Census region (%)			
Northeast	22.6	23.5	22.3*
North Central	26.1	26.1	26.1
South	35.9	33.8	36.4***
West	15.4	16.5	15.2*
Serious mental illness (%)			
Schizophrenia or other psychoses	66.1	62.9	66.9***
Bipolar disorder	54.5	71.4	50.4***
Personality disorder	10.7	23.3	7.6***
Charlson comorbidity index (%)			
0	55.2	48.5	56.9***
1–2	31.6	36.5	30.4***
3 or more	13.1	15.0	12.7***
Medicare spending (\$), mean (SD)			
Part A	8,444 (20,431)	16,088 (27,587)	6,567 (17,753)***
Part B	5,365 (7,939)	8,290 (9,584)	4,646 (7,304)***
Part D	7,626 (8,733)	8,254 (10,005)	7,472 (8,384)***

Notes. Significance refers to the difference between beneficiaries with and without any redacted Medicare Part A or Part B claims. SD, standard deviation.

* $p < .05$, ** $p < .001$, *** $p < .0001$.

study outcomes. For example, research to measure treatment effectiveness or to evaluate clinical or policy interventions designed to improve quality of care may be biased if spending data are more likely to be missing for beneficiaries with more severe forms of disease or higher Medicare expenditures.

While substance abuse redaction was unobservable at the claim level, it was possible to assess its impact at the beneficiary level by comparing claims-based spending totals for each beneficiary in our cohort to summary spending totals calculated by CMS based on the full set of unredacted claims. We found that claims-based spending totals for nearly one in five Medicare beneficiaries with serious mental illness were impacted by the redaction of substance abuse claims and that Part A expenditures for beneficiaries with redacted claims

were significantly underreported—by more than 50 percent on average. Nearly half of beneficiaries with redacted Part A spending were no longer present in the claims data, in effect making this subset of beneficiaries invisible for the purposes of claims-based cohort identification or comorbidity adjustment. Beneficiaries with and without redacted claims also differed systematically in terms of demographic characteristics and health status. Beneficiaries with redacted claims were younger, more likely to be male, had more comorbidities, and incurred higher expenditures than those without redacted claims.

The primary limitation of our analysis was that the study sample was restricted to community-dwelling beneficiaries who were continuously enrolled in fee-for-service Medicare and stand-alone Part D prescription drug plans. The findings may therefore not generalize to beneficiaries excluded from the analysis, including nursing home residents, beneficiaries enrolled in Medicare Advantage plans, and those without Part D prescription drug coverage.

As the number of individuals with substance abuse disorders entering Medicare increases, an accurate understanding of health care utilization and expenditure patterns in this population is vital to resource allocation for substance abuse prevention and treatment. SAMHSA recently issued a proposed rule that would revise the regulations governing the confidentiality of substance abuse data and restore access to Medicare and Medicaid claims data for qualified researchers (Federal Register 2016). Until the rule is finalized and implemented, or alternative changes to the regulations are made, researchers who rely on Medicare claims to analyze spending outcomes for beneficiaries with serious mental illness should be aware of the potential for substantial bias due to the nonrandom redaction of substance abuse claims. The implications of this differential redaction for research in other populations with high rates of comorbid substance abuse disorder warrant further investigation.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This was an unfunded study that was conducted as part of the lead author's doctoral dissertation research. The lead author is employed on a part-time basis by the Pharmaceutical Research and Manufacturers of America (PhRMA), which had no role in the funding or conduct of this analysis. Any views expressed here are solely those of the authors and do not reflect the views of PhRMA. Dr. Onukwugha acknowledges the receipt of grants and contracts for unrelated research from Bayer Healthcare Pharmaceuticals and Takeda Pharmaceuticals.

Disclosures: None.

Disclaimers: None.

REFERENCES

- Barosso, G. 2015. "Statistics: Impact of Redaction of Medicare and Medicaid Claims." Research Data Assistance Center (ResDAC) [accessed on February 28, 2016]. Available at <http://www.resdac.org/resconnect/articles/203>
- Buccaneer Computer Systems and Services, Inc. 2015. "Chronic Conditions Data Warehouse: Medicare Administrative Data User Guide Version 3.1" [accessed on February 28, 2016]. Available at https://www.ccwdata.org/cs/groups/public/documents/document/ccw_userguide.pdf
- Buccaneer Computer Systems and Services, Inc. 2016. "CMS Chronic Conditions Data Warehouse (CCW) Other Chronic or Potentially Disabling Condition Algorithms" [accessed on May 6, 2016]. Available at https://www.ccwdata.org/cs/groups/public/documents/document/other_cond_algos_consolidated.pdf
- Federal Register. 2016. "Confidentiality of Substance Use Disorder Patient Records; Proposed Rule." *Federal Register* 81 (26): 6988–7024 [accessed on February 28, 2016]. Available at <https://www.federalregister.gov/articles/2016/02/09/2016-01841/confidentiality-of-substance-use-disorder-patient-records>
- Frakt, A. 2015. "Addiction Research and Care Collide with Federal Privacy Rules." *New York Times: The Upshot* [accessed on February 28, 2016]. Available at http://www.nytimes.com/2015/04/28/upshot/federal-push-for-privacy-hampers-addiction-research-and-care.html?_r=0
- Frakt, A. B., and N. Bagley. 2015. "Protection or Harm? Suppressing Substance-Use Data." *New England Journal of Medicine* 372 (20): 1879–81.
- Hartz, S. M., C. N. Pato, H. Medeiros, P. Cavazos-Rehg, J. L. Sobell, J. A. Knowles, L. J. Bierut, M. T. Pato, and Genomic Psychiatry Cohort Consortium. 2014. "Comorbidity of Severe Psychotic Disorders with Measures of Substance Abuse." *Journal of the American Medical Association Psychiatry* 71 (1): 248–54.
- Rough, K., B. T. Bateman, E. Paterno, R. J. Desai, Y. Park, S. Hernandez-Dias and K. F. Huybrechts. 2016. "Suppression of Substance Abuse Claims in Medicaid Data and Rates of Diagnoses for Non-Substance Abuse Conditions." *Journal of the American Medical Association* 315 (11): 1164–6.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.