

© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12318
METHODS ARTICLE

Estimating True Resource Costs of Outpatient Care for Medicare Beneficiaries: Standardized Costs versus Medicare Payments and Charges

John T. Schousboe, Misti L. Paudel, Brent C. Taylor, Allyson M. Kats, Beth A. Virnig, Kristine E. Ensrud, and Bryan E. Dowd

Objective. To compare standardized estimates of the true resource costs of outpatient health care to the allowable and billed charges for that care among Medicare Fee for Service (FFS) beneficiaries.

Data Sources/Study Setting. Medicare Carrier and Outpatient Standard Analytic (SAF) files linked to participant data in the Study of Osteoporotic Fractures from 2004 through 2010. Participants were 3,435 female Medicare Fee for Service enrollees age 80 and older recruited in one rural and three metropolitan areas of the United States.

Study Design. We estimated standardized costs for Carrier and OP-SAF claims using Medicare payment weights, and compared them to allowable and billed charges for those claims. We used semilog linear regression to estimate the associations of age, race, bone mineral density, prior fracture, and geriatric depression scale score with allowable charges, billed charges, and standardized costs.

Results. Estimated associations of patient characteristics with standardized costs were not statistically different than the associations with allowable charges (chi-squared [χ^2]: 8.6, $p = .13$) but were different from associations with billed charges (χ^2 : 25.5, $p < .001$).

Conclusion. Allowable charges for outpatient utilization in the Carrier file and OP-SAF may be good surrogates for standardized costs that reflect patient medical and surgical acuity.

Key Words. Outpatient costs, Medicare payment, Carrier file, outpatient SAF file, health care costs

Accurately measuring health care costs and understanding the determinants of those costs are essential to improving health care efficiency. Researchers use Medicare payments as a surrogate measure of health care costs for three reasons: (1) those age 65 and older have the highest per person health care

costs; (2) Medicare is the insurance carrier that is the primary payer for 95 percent of this segment of the population, allowing for such studies to be representative of all regions of the United States; and (3) Medicare strives to pay a specific amount needed to care for that condition based on the resource intensity needed for patients with that condition.

While health care structure and local medical practice variation are important drivers of health care costs (Wennberg et al. 1989; Rosenthal et al. 1997; Fisher and Wennberg 2003; Fisher et al. 2003a,b), individual patient characteristics can also be important (Zuckerman et al. 2010; Reschovsky, Hadley, and Romano 2013) but are not always available in a study limited to administrative claims data. Linking large observational cohort studies, such as the Study of Osteoporotic Fractures (SOF) and Framingham, with Medicare claims may improve our understanding of how individual patient characteristics influence health care costs because these studies have usually recorded phenotypic characteristics of participants with greater detail and accuracy than what is available in claims or medical record data.

Using datasets created by linking observational cohort to Medicare claims for these purposes, however, can be challenging. First, Medicare adjusts payments to both hospital and outpatient providers for local health care input prices (labor and capital), which can result in regional variability in Medicare payments. Even large cohort studies recruit participants from only a handful of geographic locations, and Medicare payment data from these locations may not be generalizable to the general U.S. population. Second, Medicare pays some facilities that provide inpatient and/or outpatient services differently according to the degree to which they employ physicians in training, the proportion of indigent clients under their care, and other characteristics. These adjustments to Medicare payments may not be relevant to the determination of associations between individual characteristics and resource costs required for health care. For example, whether or not a patient received care in an area that has low labor and capital costs conceptually would not be relevant to the association of a specific participant characteristic (such as

Address correspondence to John T. Schousboe, M.D., Ph.D., Health Research Center, Park Nicollet Institute for Research and Education, 3800 Park Nicollet Blvd, Minneapolis, MN 55416; e-mail: schouj@parknicollet.com. Misti L. Paudel, Ph.D., M.P.H., is with the NORC, University of Chicago, Bethesda, MD. Brent C. Taylor, Ph.D., M.P.H., and Kristine E. Ensrud, M.D., M.P.H., are with the Center for Chronic Disease Outcomes Research, Minneapolis VAMC, Minneapolis, MN. Allyson M. Kats, M.S., is with the Chronic Disease Research Group, Minneapolis Medical Research Foundation, Minneapolis, MN. Beth A. Virnig, Ph.D., and Bryan E. Dowd, Ph.D., are with the Division of Health Policy and Management, University of Minnesota, Minneapolis, MN.

chronic kidney disease) with the resources required to care for patients with that condition. Removing these adjustments from Medicare payments may yield standardized cost estimates, for example, dollar-weighted estimates of the costs required to care for the patient, where each unit of utilization has been multiplied by the same unit price regardless of geographic location or the types of patients treated at the facility.

We have previously shown that standardized cost estimates for acute hospital stays are significantly different from actual payments to hospitals due to the adjustments Medicare makes to hospital provider payments described above, even though both are based on DRG payment weights (Schousboe et al. 2014). In this paper, we extend these analyses to payments for both facility-based and nonfacility-based outpatient services. Similar to inpatient costs, Medicare bases both provider payments and standardized cost estimates for outpatient on payments weights and assigns these weights to units of utilization based on Healthcare Common Procedure Coding System (HCPCS) codes for nonfacility outpatient services, and Ambulatory Payment Classifications (APC) codes for facility-based outpatient services. Our primary research question was whether or not the adjustments Medicare makes to provider payments for local geographic input prices and (in the case of facility-based services) other facility characteristics result in significant differences between outpatient costs calculated from standardized cost estimates, provider payments, and billed charges for these services.

As we did for inpatient hospital costs, we used Medicare claims data for participants enrolled in both the SOF and Medicare Fee for Service during part or all of the time period January 1, 2004, through December 31, 2010, to estimate the change in rank order of estimated yearly outpatient costs among SOF participants when using standardized costs instead of provider payments or instead of billed charges.

METHODS

The Study of Osteoporotic Fractures recruited 9,704 Caucasian women age 65 or older between 1986 and 1988 in four geographic regions of the United States: Baltimore, MD; Minneapolis, MN; Portland, OR; and a rural area (Monongahela Valley) near Pittsburgh, PA (Cummings et al. 1990). At the sixth SOF study visit between 1996 and 1998, 662 African American women age 65 and older were also recruited (Cauley et al.

2005). Using validated methods detailed in previous publications (Schousboe et al. 2013, 2014), we successfully matched Medicare claims to 92 percent ($n = 9,228$) of surviving women still enrolled in SOF as of January 1, 1991, when outpatient Medicare claims first became available. Because the Medicare files required to calculate standardized outpatient costs (described in the Appendix) are available only from 2004 onward, we used the denominator file to identify surviving SOF participants still enrolled in Medicare Fee for Service parts A and B for all or part of calendar years 2004 through 2010 inclusive.

We used two Medicare claims files to identify items of ambulatory care. The *Carrier file* contains claims from providers such as physicians and some outpatient claims of utilization occurring in facilities such as hospitals, skilled nursing facilities, and hospices. The *Outpatient Standard Analytic File (OP-SAF)* contains claims for certain types of utilization in facilities paid under a different system, the Hospital Outpatient Prospective Payment System.

We characterized provider payments as allowable charges, which is the sum of what Medicare pays providers directly plus the deductible and copayments for which patients are responsible. The allowable charges are listed for each unit of utilization as a separate variable in the Carrier file, and in the OP-SAF file they are calculated as the sum of the Medicare payment and the beneficiary responsibility amount for each item of utilization. We identified billed charges as the submitted charge in the Carrier file and as the claim total charge in the OP-SAF.

We calculated standardized costs by applying payment weights (unadjusted for local geographic variation in labor and capital input prices) assigned by Medicare to HCPCS codes (for the Carrier file) and APC codes (for the OP-SAF). The unadjusted payment weights reflect the resource intensity (standardized for the entire United States) required to deliver the service. The methods used to calculate standardized costs for all outpatient services are fully described in the Appendix.

Individual Characteristics as Predictors of Costs

We tested the associations of age, race, femoral neck bone mineral density (BMD), prior clinical fracture, and depression with standardized costs, allowable charges, and billed charges, hypothesizing that each of these would be associated with health care costs and utilization on account of their association with comorbidity burden and health status. These covariates were assessed at

the eighth SOF study visits at the four study sites between January 2002 and April 2004, at which 4,388 surviving participants attended.

We measured femoral neck BMD on Hologic 4500A densitometers (Hologic Inc, Bedford, MA, USA). SOF cohort study staff assessed self-reported depression with the 15-item version of the Geriatric Depression Scale (GDS). At study entry, participants were asked if they had any fractures since age 50. After study entry, we identified incident fractures by follow-up contact with participants or proxies every 4 months by mailed postcards or phone, with a 95 percent response rate through 2010. SOF investigators confirmed incident fractures by physician adjudication of radiographic reports. We categorized women as having had a prior fracture as of the eighth SOF visit if they self-reported a prior fracture at the baseline visit or if they had a radiographically confirmed incident clinical fracture between the baseline and eighth visits.

Statistical Analysis

For each calendar year, we calculated the sums of standardized costs, allowable charges, and billed charges (for items of utilization for which payment was not denied and a standardized cost could be established) in the Carrier file and OP-SAF separately, and for both added together (representing total outpatient utilization). We estimated Spearman correlations between these with the individual participant as the unit of analysis. We expressed the distribution of each of these nine cost sums as quintiles and calculated the change in quintile rank order when using standardized costs rather than allowable or billed charges as the standardized cost quintile number minus the allowable or billed charge quintile number.

For each of the nine cost sums, we regressed the log of the costs for the first full calendar year after the eighth SOF visit was concluded (2005) on age (as of January 1, 2005), race, prior clinical fracture, femoral neck BMD, and GDS score with linear regression models using *Stata version 13.0* (STATA Corp., College Station, Texas, USA). We tested each model for heteroskedasticity and for mis-specification with the Ramsey powers test and Pregibon's linktest (Pregibon 1980). We tested the hypothesis that the parameter estimates of association between predictors and standardized costs were the same as with allowable or with billed charges for total outpatient utilization (Carrier file plus OP-SAF utilization) and separately for Carrier costs and OP-SAF costs alone for 2005 with the *suest* command of *Stata* (a modified Hausman test).

Sensitivity Analyses

Among all SOF participants enrolled in Medicare fee for service for at least 1 month from 2004 through 2010 inclusive, there were one or more items of utilization for more than 75 percent of all U.S. geographic locations defined by carrier number and locality code. However, half of all items of utilization were in one of three geographic locations (Minnesota, rural Pennsylvania, and Baltimore), and the geographic adjustment factors for these three locations were within 8 percent of one another. To examine if a much wider spread of geographic adjustment among these three sites would alter our results, we repeated the analyses described above for total outpatient utilization of a hypothetical population with the same individual characteristics, but reducing allowable charges for rural Pennsylvania to correspond to what they would be if rural Pennsylvania participants received their care in Oklahoma City (the least expensive large urban geographic location in 2005) and by raising allowable charges for Baltimore and surrounding counties to correspond to what they would be if Baltimore participants received their care in San Francisco (the most expensive large city in 2005).

Because claims for some medical services delivered during an inpatient stay are found in the Carrier file (such as many physician and surgeon services), we also repeated all of the previously described analyses after excluding items of utilization that occurred during an inpatient stay. We determined the dates of inpatient stays from the Medical Provider and Analysis Review (MedPAR) file.

RESULTS

From January 1, 2004, through December 31, 2010, 3,535 SOF participants enrolled in Medicare Fee for Service Parts A and B for one or more of these seven calendar years; 3,106 had claims in both the Carrier and Outpatient SAF files, 219 had only Carrier file claims, and 210 had only Outpatient SAF claims. In each of the 7 years, 90–92 percent of individuals had a paid Carrier claim, and 78–80 percent had a paid OP-SAF claim. Over that time span, there were 656,322 paid Carrier claims, and we were able to establish standardized costs for 609,906 (92.9 percent). Similarly, there were 367,815 paid OP-SAF claims, and we were able to establish standardized costs for 313,473 (85.2 percent).

The medians and interquartile ranges of the nine cost sums for calendar year 2005 are shown in Table 1, and for all years in the Table S1a, b, and c.

Carrier standardized costs and provider payments were more than twice that of OP-SAF standardized costs and allowable charges. Billed charges were more than twice allowable charges and standardized costs. Standardized costs were particularly highly correlated with allowable charges and slightly less with billed charges (Table 1).

When total outpatient costs for calendar year 2005 are expressed as allowable charges rather than standardized costs, 12 percent of individuals change rank order quintile, whereas 25 percent change rank order quintile when costs are expressed as billed charges rather than standardized costs (Figure 1). Nearly identical results are seen for all of the other six calendar years (Table S2a, b, and c).

When we regressed the log of total outpatient costs on age, race, femoral neck BMD, prior clinical fracture, and GDS score, each of the parameter coefficients and their 95 percent confidence intervals were very close, with the possible exception of race, whether costs are expressed as standardized costs or allowable charges (Table 2). In both regressions, age, GDS score, and prior clinical fracture were significantly associated with the log of total outpatient costs, whereas race and femoral neck BMD were not. Considering all predictors together, we did not reject the hypothesis that the predictor parameter estimates of association with standardized carrier costs were the same as estimates of association with allowable carrier charges (chi-squared [χ^2]: 8.57, $p = .13$).

In contrast, the associations of predictor variables with total billed charges were significantly different than their associations with total outpatient standardized costs (χ^2 : 25.5, $p < .001$, Table 3). Femoral neck BMD was not significantly associated with the log of total outpatient standardized costs, but it was with the log of total outpatient billed charges.

When Carrier file and OP-SAF claims are considered separately, we found that the associations of predictor variables with Carrier standardized costs and with allowable charges were nearly identical (χ^2 : 4.72, $p = .45$, Table S1a). In contrast, we found that the associations of these predictors with standardized Carrier costs were substantially different than with Carrier billed charges (χ^2 : 38.4, $p < .001$, Table S1b). For OP-SAF costs, there were small differences in the associations of standardized costs, allowable charges, and billed charges with these same predictors (Tables S2a and b). Considering all predictors together, we did not reject the hypotheses that the parameter estimates of association with standard costs and were the same as with allowed charges (χ^2 : 7.52, $p = .18$, Table S2a) or with billed charges (χ^2 : 8.74, $p = .12$, Table S2b).

Table 1: Distributions of Standardized Costs, Allowable Charges, and Billed Charges for 2005

	<i>Carrier File Utilization</i> (<i>n</i> = 2,623) [†]	<i>OP-SAF Utilization</i> (<i>n</i> = 2,235) [†]	<i>Total Outpatient Utilization*</i> (<i>n</i> = 2,847) [†]
Standardized costs, median (IQR)	\$1,683 (693–3,353)	\$681 (225–1,725)	\$2,243 (882–4,469)
Allowable charges, median (IQR)	\$1,696 (711–3,353)	\$722 (226–1,822)	\$2,248 (871–4,473)
Billed charges, median (IQR)	\$3,661 (1,422–7,619)	\$981 (104–3,567)	\$4,963 (1,869–10,462)
Difference between standardized cost and allowable charges, median (IQR)	–\$9 (–69 to 41)	–\$31 (–232 to 47)	\$23 (–92 to 223)
Spearman correlation between standardized cost and allowable charges	0.99	0.96	0.99
Difference between standardized cost and billed charges, median (IQR)	–\$1,816 (–4,228 to –642)	–\$437 (–2,066 to 0)	–\$2,552 (–5,894 to –873)
Spearman correlation between standardized cost and billed charges, median (IQR)	0.98	0.91	0.96

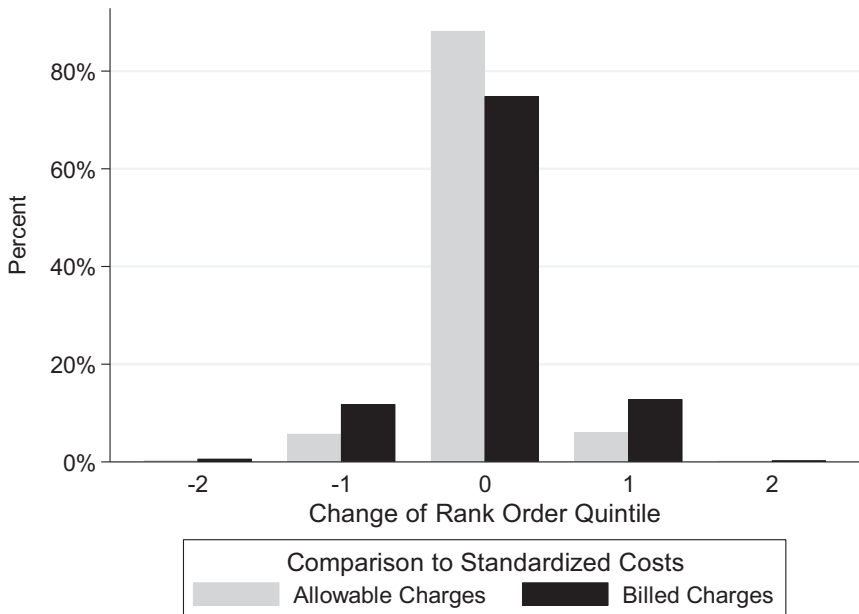
*Carrier file plus OP-SAF utilization.

†Number in each column are numbers of women with one or more paid claims for which standardized costs could be estimated.

Sensitivity Analysis

For a hypothetical study population with the same individual characteristics, but with allowable charges reduced by 4 percent for rural Pennsylvania and increased by 18 percent in Baltimore (to gage the effect of having one cohort study site recruited from a very inexpensive large urban geographic location and another from a very expensive large urban location), the associations of the predictors with all standardized costs were different than with allowable charges (χ^2 : 13.03, p = .02, considering all predictors together). However, the conclusions remained the same (GDS score and prior fracture are associated

Figure 1: Shift of Rank Order Quintile When Using Standardized Costs of Total Outpatient Utilization Instead of Allowed Charges or Billed Charges



with costs; and age, race, and BMD are not) when we used standardized costs rather than allowable charges in this hypothetical example.

All of the results described above were unchanged when we excluded line items of utilization that occurred during inpatient stays (20.5 percent of all items) (data not shown).

DISCUSSION

In this study, we analyzed the differences between standardized costs, allowable charges, and billed charges using data from SOF participants merged with their Fee for Service Medicare claims. We found that standardized resource cost estimates for total outpatient (the sum of Carrier and OP-SAF) claims are reasonably close to allowable charges, whereas the differences between standardized costs and billed charges are greater. We did not find consistent differences in the associations of predictor variables with allowable charges compared to standardized costs, but they did demonstrate significant

Table 2: Associations of Predictors with Standardized Total Outpatient Costs and Allowable Charges for 2005 ($n = 1,531$)

Predictor	Mean [SD] or N (%)	Parameter Estimates (95% CI)		Parameter Difference, p -value*
		Log Standardized Costs	Log Provider Payments	
Age	85.9 years (3.6)	-0.095[†] (-0.168 to -0.021)	-0.092[†] (-0.164 to -0.019)	.65
Race	Black: 191 (12.4%) White: 1,346 (87.6%)	0.063 [‡] (-0.098 to 0.225)	0.094 [‡] (-0.067 to 0.254)	.02
Femoral neck BMD	0.641 g/cm ² (0.124)	0.035 [§] (-0.018 to 0.087)	0.039 [§] (-0.013 to 0.091)	.22
GDS score (0-15)	2.25 (2.59)	0.072[§] (0.017-0.127)	0.068[§] (0.014-0.123)	.45
Prior clinical fracture	No: 698 (45.6%) Yes: 833 (54.4%)	0.168[¶] (0.066-0.270)	0.168[¶] (0.066-0.269)	.83

Note. Parameter coefficients with a p -value of association < 0.05 are in bold.

*Wald Chi test of hypothesis that the parameter coefficients for that predictor or the same; considering all parameters together, $\chi^2 = 8.57$, p -value = .13.

[†]Number log(costs) standard deviation changes per 5 years increase in age.

[‡]Number log(costs) standard deviation changes black compared to white.

[§]Number of log(costs) standard deviation changes per standard deviation increase in predictor.

[¶]Number of log(costs) standard deviation changes prior fracture compared to no prior fracture.

Table 3: Associations of Predictors with Total Outpatient Standardized Costs and Billed Charges for 2005 ($n = 1,531$)

Predictor	Mean [SD] or N (%)	Parameter Estimates (95% CI)		Parameter Difference, <i>p</i> -value*
		Log Standardized Costs	Log Billed Charges	
Age	85.9 years (3.6)	-0.095[†] (-0.168 to -0.021)	-0.101 (-0.174 to -0.028)	.28
Race	Black: 191 (12.4%) White: 1,346 (87.6%)	0.063 [‡] (-0.098 to 0.225)	0.115 (-0.046 to 0.275)	.01
Femoral neck BMD	0.641 g/cm ² (0.124)	0.035 [§] (-0.018 to 0.087)	0.053 (0.01-0.105)	.01
GDS score (0-15)	2.25 (2.59)	0.072[§] (0.017-0.127)	0.079 (0.024-0.134)	.14
Prior clinical fracture	No: 698 (45.6%) Yes: 833 (54.4%)	0.168[¶] (0.066-0.270)	0.170 (0.069-0.272)	.44

Note. Parameter coefficients with a *p*-value of association <0.05 are in bold.

*Wald Chi test of hypothesis that the parameter coefficients for that predictor or the same; considering all parameters together, $\chi^2 = 25.5, p < .001$.

[†]Number log(costs) standard deviation changes per 5 years increase in age.

[‡]Number log(costs) standard deviation changes black compared to white.

[§]Number of log(costs) standard deviation changes per standard deviation increase in predictor.

[¶]Number of log(costs) standard deviation changes prior fracture compared to no prior fracture.

differences in the associations with billed charges compared to standardized costs. These findings support previous statements that billed charges are a poor proxy for standardized costs (Finkler 1982; Taira et al. 2003; Polsky and Glick 2009). For research questions where these local geographic factors are not relevant, modeling costs as allowable charges would have *the potential* to introduce stochastic noise that would reduce statistical power or bias estimated associations of predictors with costs. Our data indicate that the adjustments to Medicare outpatient allowable charges for local geographic and other provider characteristics are modest enough that this is not a major concern, especially for Carrier file utilization.

The potential for biased estimates of association of predictors with costs is low, but not zero, when modeling costs as allowable charges rather than standardized costs. In particular, we advise caution in studies that recruit large proportions of their participants from locations with geographic adjustment factor values at the extremes of the distribution (such as a very expensive location like San Francisco and an inexpensive location like Oklahoma City). In these instances, it is possible that meaningful differences between standardized resource costs and allowable charges would be more apparent, and sensitivity analyses with a subset of claims for which standardized costs can be established may be appropriate.

We believe that calculating allowable charges from Medicare payment and beneficiary responsibility variables in the outpatient files is easier than having to estimate standard costs for large numbers of outpatient claims from relative value unit (RVU) and other standard payment files. Researchers using the latter approach have to merge those standard payment files to actual claims data. Moreover, RVU values and relative payment weights, respectively, for Carrier and OP-SAF claims are not readily available for a significant minority of claims. We believe that our results are generalizable to the majority of studies estimating the associations of individual characteristics with outpatient health costs done within the United States, whether the sources of those characteristics are prospective cohort studies, electronic health records, or administrative data.

Our study extends the findings of O'Donnell et al. (2013), who estimated the correlations of aggregated standardized costs and allowable charges within hospital referral regions for all Medicare Fee for Service patients experiencing an acute myocardial infarction in 2007. They calculated standardized costs and allowable charges using methods very similar to ours. They too found substantially greater variation between Medicare payments and standardized hospital costs for inpatient stays, and relatively minor differences

between standardized costs and actual payments for outpatient claims. However, our study used the individual patient as the unit of analysis and also examined facility outpatient utilization payments (in the OP-SAF) separately from nonfacility outpatient utilization payments (in the Carrier file).

Our study has important strengths. Recruitment into the SOF used population-based listings such as driver's license registries, and participants' characteristics (matched for age and race) are very similar to those of the U.S. population. The linkage of Medicare claims with SOF cohort data allowed us to specifically test hypotheses that predictor parameter coefficients would not be significantly different if costs were modeled as standardized costs rather than allowable charges or billed charges.

We also recognize that our study has several limitations. Our study population is very old (86.7 [SD: 4.1] years) because we could not calculate standardized costs for utilization before 2004. We recruited our study population from only four geographic regions of the United States and did not include some ethnic groups, specifically Hispanic and Asian Americans. We assumed that RVU and other payment weights truly capture the dollar-weighted resource costs required to deliver an outpatient service. Our results are not applicable to private health care insurer payments to providers for those under age 65.

In conclusion, we found that outpatient allowable Medicare charges in the Carrier file and OP-SAF are a reasonable approximation of standardized costs. In contrast, we also found that billed charges are sufficiently different from standardized costs that use of billed charges as a proxy for standardized costs may lead to biased estimates of association of predictors with outpatient costs.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This study was done primarily under funding from the National Institute for Aging/National Institutes of Health, primarily grant number R01 AG038415-01. The Study of Osteoporotic Fractures is also supported by the National Institute for Aging under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576.

Disclosures: None.

Disclaimers: None.

REFERENCES

- Cauley, J. A., L. Y. Lui, K. E. Ensrud, J. M. Zmuda, K. L. Stone, M. C. Hochberg, and S. R. Cummings. 2005. "Bone Mineral Density and the Risk of Incident Nonspinal Fractures in Black and White Women." *Journal of the American Medical Association* 293 (17): 2102–8.
- Cummings, S. R., D. M. Black, M. C. Nevitt, W. S. Browner, J. A. Cauley, H. K. Genant, S. R. Mascioli, J. C. Scott, D. G. Seeley, P. Steiger, and T. M. Vogt 1990. "Appendicular Bone Density and Age Predict Hip Fracture in Women. The Study of Osteoporotic Fractures Research Group." *Journal of the American Medical Association* 263 (5): 665–8.
- Finkler, S. A. 1982. "The Distinction between Cost and Charges." *Annals of Internal Medicine* 96 (1): 102–9.
- Fisher, E. S., and J. E. Wennberg. 2003. "Health Care Quality, Geographic Variations, and the Challenge of Supply-Sensitive Care." *Perspectives in Biology and Medicine* 46 (1): 69–79.
- Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas, and E. L. Pinder. 2003a. "The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care." *Annals of Internal Medicine* 138 (4): 273–87.
- . 2003b. "The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care." *Annals of Internal Medicine* 138 (4): 288–98.
- O'Donnell, B. E., K. M. Schneider, J. M. Brooks, G. Lessman, J. Wilwert, E. Cook, G. Martens, K. Wright, and E. A. Chrischilles. 2013. "Standardizing Medicare Payment Information to Support Examining Geographic Variation in Costs." *Medicare Medicaid Research Review* 3 (3): E1–20.
- Polsky, D., and H. Glick. 2009. "Costing and Cost Analysis in Randomized Controlled Trials: Caveat Emptor." *Pharmacoeconomics* 27 (3): 179–88.
- Pregibon, D. 1980. "Goodness of Link Tests for Generalized Linear Models." *Applied Statistics* 29: 15–24.
- Reschovsky, J. D., J. Hadley, and P. S. Romano. 2013. "Geographic Variation in Fee-for-Service Medicare Beneficiaries' Medical Costs Is Largely Explained by Disease Burden." *Medical Care Research and Review: MCRR* 70 (5): 542–63.
- Rosenthal, G. E., D. L. Harper, A. Shah, and K. E. Covinsky. 1997. "A Regional Evaluation of Variation in Low-Severity Hospital Admissions." *Journal of General Internal Medicine* 12 (7): 416–22.
- Schousboe, J. T., M. L. Paudel, B. C. Taylor, B. A. Virnig, J. A. Cauley, J. R. Curtis, and K. E. Ensrud. 2013. "Magnitude and Consequences of Misclassification of Incident Hip Fractures in Large Cohort Studies: The Study of Osteoporotic Fractures and Medicare Claims Data." *Osteoporosis International* 24 (3): 801–10.
- Schousboe, J. T., M. L. Paudel, B. C. Taylor, L. W. Mau, B. A. Virnig, K. E. Ensrud, and B. E. Dowd. 2014. "Estimation of Standardized Hospital Costs from Medicare Claims That Reflect Resource Requirements for Care: Impact for Cohort Studies Linked to Medicare Claims." *Health Services Research* 49 (3): 929–49.

- Taira, D. A., T. B. Seto, R. Siegrist, R. Cosgrove, R. Berezin, and D. J. Cohen. 2003. "Comparison of Analytic Approaches for the Economic Evaluation of New Technologies alongside Multicenter Clinical Trials." *American Heart Journal* 145 (3): 452–8.
- Wennberg, J. E., J. L. Freeman, R. M. Shelton, and T. A. Bubolz. 1989. "Hospital Use and Mortality among Medicare Beneficiaries in Boston and New Haven." *New England Journal of Medicine* 321 (17): 1168–73.
- Zuckerman, S., T. Waidmann, R. Berenson, and J. Hadley. 2010. "Clarifying Sources of Geographic Differences in Medicare Spending." *New England Journal of Medicine* 363 (1): 54–62.

SUPPORTING INFORMATION

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

Appendix SA1: Author Matrix.

Appendix SA2. Methods for Estimating Standardized Costs.