Assessing the RUG-III Resident Classification System for Skilled Nursing Facilities

Chapin White, M.P.P., Steven D. Pizer, Ph.D., and Alan J. White, Ph.D.

Resource utilization groups, version III (RUG-III) is used by CMS to classify skilled nursing facility (SNF) residents into Medicare payment groups. Using a sample of 1,304 SNF residents with Medicare-covered stays, we find that RUG-III only explains 10.4 percent of the variance in total per diem costs. RUG-III explains variance in staff-time costs fairly well, but does not explain variance in non-therapy ancillary costs. Receipt of special treatments such as intravenous medications and respiratory therapy is strongly associated with high residual costs (p<0.01). Modifications to the RUG-III system can increase its variance explanation.

INTRODUCTION

As mandated by the Balanced Budget Act of 1997, the Health Care Financing Administration on July 1, 1998 began phasing in a prospective payment system (PPS) for Medicare-covered stays in SNFs. The prospective payment is meant to cover nearly all costs of providing care, including routine service, ancillary and capital-related costs, with only a few exceptions such as physician services. The Balanced Budget Act of 1997 requires that payments be adjusted for case mix using a "resident classification system, established by the Secretary [of Health and Human Services], that accounts for the relative resource utilization of different patient types"(U.S. Congress, 1997). RUG-III is the classification system currently used by CMS to classify residents for Medicare payment. Our objectives were to test the variance explanation of the RUG-III resident classification system, identify reasons for its low variance explanation, and test the variance explanation of modified versions of RUG-III.

RUG-III was developed in the early 1990s using staff-time data from a sample of nursing facilities in six States (Fries et al., 1994). Classification is based on residents' physical functioning, disease diagnoses, health conditions, and treatments received, as recorded in the minimum data set, version 2.0 (MDS-2.0), a lengthy and detailed data collection instrument filled out by SNF staff. RUG-III first tests whether a SNF resident qualifies for each of the seven major categories: (1) rehabilitation, (2) extensive services, (3) special care, (4) clinically complex, (5) impaired cognition, (6) behavior problems, and (7) reduced physical function. Residents who qualify for a major category are then split into mutually exclusive groups, with names such as "Very High Rehabilitation A" and "Extensive Services 3" (*Federal Register*, 1998). A resident may meet the qualification criteria for many RUG-III groups, but a resident is only assigned to the group with the highest payment rate. Between January and March

7

2001, 77.4 percent of SNF residents were assigned to the rehabilitation major category, and 13.3 percent were assigned to the extensive services major category (U.S. General Accounting Office, 2002). Within the rehabilitation major category, assignment to a group depends on the number of minutes of rehabilitation therapy received per week and the resident's limitations in activities of daily living. Within extensive services, assignment to a group depends on intravenous medications, parenteral feeding, and whether the resident qualifies for other major categories.

One of the fundamental tradeoffs in the payment of health care providers is efficiency in production versus selection and skimping (Newhouse, 1996). By setting payment rates prospectively, payers encourage productive efficiency, meaning minimizing the cost of producing a given service. This desirable response to prospective payment comes at the cost of creating incentives to reduce the level and quality of services provided (skimping) and to seek low-cost patients and shun high-cost patients (selection). By classifying individuals into case-mix groups, the incentive for providers to select low-cost residents can be lessened. The effectiveness of a case-mix classification selection depends on its ability to sort individuals into groups with homogenous resource needs. This effectiveness is typically measured using the R^2 statistic from a regression of costs on case-mix group dummy variables. The R^2 statistic is equivalent to the percent of variance in costs explained by the case-mix classification system. In assessing a case-mix classification system, there is no clear cutoff separating an acceptably high R^2 from an unacceptably low R^2. In assessing RUG-III, however, we should expect the R^2 to be high because classification is largely determined by the services provided and costs incurred rather than diagnoses. Furthermore,

because the SNF PPS has no outlier provision, the standards for an acceptable R^2 should be higher than in other settings in which there is an outlier provision (such as short-stay hospitals or home health).

We use resident-level data on clinical characteristics and the costs of providing care to estimate the effectiveness of the RUG-III case-mix classification system. We find that the RUG-III case-mix groups explain little of the variance in total costs, and that certain resident characteristics (such as receiving intravenous medications) are strongly associated with having high costs relative to other individuals in the same case-mix group. We test simple modifications of the RUG-III system, and find that they achieve higher variance explanation.

METHODS

Data Sources

For all of our analyses we use data from the CMS-funded SNF staff-time measurement (STM) study which took place in 1995 and 1997. The STM study recorded the number of minutes per day of nurse and rehabilitation therapist staff time required to care for SNF residents. The STM included data on 3,933 Medicare, Medicaid and self-pay residents in 150 Medicare-certified SNF units in 12 States (Kansas, Maine, Mississippi, Ohio, South Dakota, Texas, Washington, California, Florida, Maryland, Colorado, and New York). To collect staff-time data, all nursing staff used electronic wands over a period of 48 hours to record direct resident care time. Physical, occupational, and speech therapy time was collected over the span of a week. Non-resident specific nursing time (such as meetings, administration, breaks, and unit maintenance) was allocated equally across all unit residents. The STM data for all 3,933 residents were merged with

MDS-2.0 assessments that were filled out near the time of the STM. The MDS-2.0 assessment data were checked for missing, invalid, and out-of-range responses and were recoded when the intention was obvious (such as replacing—with 0. Dichotomous variables were created based on certain non-dichotomous variables. For example, G1a_not ("transfer did not occur during last 7 days") was created based on G1aa ("transfer activities of daily living self-performance") and O3_any ("1 or more injections received during the last 7 days") was created based on O3 ("number of days injections of any type received during the last 7 days").

This research differs from previous SNF case-mix research in that our data set includes Medicare Part A and Part B claims data on ancillary services such as diagnostic services, supplies, and prescription drugs. Of the 3,933 records in the STM file, 1,344 had a Medicare SNF claim that met the following criteria for a Medicare stay: the provider identifier in the Medicare claim(s) matched that of the facility surveyed in the STM and the Medicare claim(s) covered a date within 15 days of the STM study week. These stays were defined as Medicare stays meaning that we assumed the SNF care on the day of the STM was paid for by Medicare. A relatively large fraction of stays, 1,344 out of 3,933 or 34 percent, were Medicare stays, which reflects the fact that high-acuity facilities were chosen to participate in the STM. The process of building the data set is described in more detail by White, Pizer, and White (1999).

Estimating Total Per Diem Costs

The STM data report the number of minutes of resident care provided by nursing and therapy staff. Total per diem staff-time costs for each resident were estimated by multiplying the number of occupation-specific minutes for that resident by that occupation's salary and summing across all of the occupations (White, Pizer, and White, 1999) . We found that fringe benefits represent 24 percent of the total compensation for nursing home employees (U.S. Department of Labor, 1999). Staff time costs were, therefore, inflated by a factor of 1.3 to account for these fringe benefits.

The covered charges reported in the claims are routinely discounted by the fiscal intermediary, on the basis of audited reasonable cost. We calculated a cost-to-charge discount factor (the ratio of total Part A allowed cost to total Part A charges) for each facility based on Medicare cost reports. This discount factor was applied to the facility's non-therapy ancillary charges. For facilities with missing cost report data, discount factors were imputed using the average discount factor among facilities from the same State, year, and setting (hospital-based versus freestanding). To account for overhead costs such as capital, building maintenance, and non-clinical staff, $55.88 was added to the estimated per diem costs for each resident—this equals the overhead component of the Federal RUG-III payment rates in 1998 (*Federal Register*, 1998). Total estimated per diem cost (hereafter referred to simply as "total costs") for each resident equals staff time costs estimated from the STM (including nursing and therapy) plus non-therapy ancillary costs estimated from Medicare claims plus overhead costs estimated from the Federal reimbursement rate. Costs reflect nominal dollars (note, though, that the same wage weights were applied to the 1995 and 1997 staff-time minutes and that the wages were estimated using 1996 data).

Creation of the Analytic Data Set

For all 3,933 observations, the STM data were merged with the MDS-2.0 assessment data. For the 1,344 observations with a valid Medicare SNF stay claim, the STM and MDS-2.0 data were merged with Medicare claims data. Observations for 38 Medicare stays and 102 non-Medicare stays had missing or invalid MDS-2.0 data and were, therefore, excluded from the sample. Two Medicare stay observations from the same facility were excluded from analysis because of their extremely unusual cost patterns. This yielded a final analytic data set with 1,304 Medicare-stay observations and 2,487 non-Medicare observations. The RUG-III grouper was used, with individuals assigned to the highest-payment group using payment rates from the *Federal Register* (1998) and the SAS Code for RUG III, version 5.12, (Health Care Financing Administration, 2000).

Analytic Strategy

All analyses were performed using SAS/STAT version 6.12. We first described the cost variables for Medicare stays using univariate analysis. We then tested the variance explanation of RUG-III for different measures of costs and with different samples (Medicare stays only and all stays). Variance explanation was measured using the R^2 statistic from a regression of costs on a set of dummy variables representing the RUG-III groups.

We then tested whether certain clinical characteristics were associated with having high costs within RUG-III groups. We first calculated mean costs within each RUG-III group and then calculated residual cost, equal to the difference between an individual's cost and the mean costs within that individual's RUG-III group. This residual cost measure is equivalent to the residual from a regression of costs on a set of dummy variables representing the 44 RUG-III groups. *T*-tests were then performed using all (362) dichotomous MDS-2.0 items tested one at a time as predictors of residual costs. An additional *t*-test was also performed on a subsample of residents assigned to the rehabilitation category, using qualification for the extensive services category as a predictor of total costs. To measure *t*-test *p*-values, we assumed unequal variance across classes. We measured 95 percent confidence intervals for the within-class means using the standard errors generated by the SAS/STAT *t*-test procedure.

One possible modification of RUG-III is to replace the hierarchical classification system with a non-hierarchical system, in which a resident may be assigned to more than one group. We tested two non-hierarchical systems. In the first, the 44 mutually exclusive RUG-III groups are replaced by 44 dummy variables, each set to 1 or 0 depending on whether the resident met the qualification criteria for a particular RUG-III group. This non-hierarchical version differs from the standard RUG-III in that the groups are not mutually exclusive. A regression model was run of total costs on these 44 dummy variables. A second non-hierarchical version of RUG-III (the "simple" non-hierarchical model) was also tested. This version includes four dummy variables based on whether the resident met the qualification criteria for a group or groups in the following major categories: rehabilitation, extensive services, special care domain, and clinically complex.

RESULTS

Mean total per diem costs are $234.56. The distributions of total per diem costs, staff time costs, and non-therapy ancillary costs are summarized in Table 1. Mean per

Table 1

Univariate Analysis of Per Diem Costs Among Medicare Stays[1]

Item	Total Per Diem[2]	Staff Time	Non-Therapy Ancillary
	$234.56	$104.66	$74.02
Mean[3]	($226.89-$242.23)	($102.10-$107.22)	($67.36-$80.68)
Median	196.34	99.31	28.63
Minimum	72.86	13.43	0.00
Maximum	1,993.53	409.88	1,841.71
Standard Deviation	141.34	47.21	122.75
Skewness	3.83	0.85	4.88

[1] N=1,304.

[2] Total costs include staff time costs, non-therapy ancillary costs, and a flat $55.88 overhead amount for all observations. This overhead is the amount of the "non-case mix" component of the Federal per diem rates (*Federal Register*, 1998).

[3] Ninety-five percent confidence intervals are shown in parentheses.

SOURCES: White, C., Harvard University, Pizer, S.D., Boston University and Department of Veterans Affairs, and White, A.J., Abt Associates, 2002.

Table 2

Variance Explanation of the RUG-III Classification System

Sample	Cost Measure	Number of Observations	Mean Cost	Percent of Variance Explained
Medicare Stays	Total	1,304	$234.56	10.4
Medicare Stays	Staff Time	1,304	$104.66	20.6
Medicare Stays	Non-Therapy Ancillary	1,304	$74.02	7.2
All Stays	Staff Time[1]	3,791	$75.56	40.1

[1] Only staff time costs are available for non-Medicare stays. Non-Medicare stays are defined as stays among residents who could not be matched to Medicare Part A or Part B claims data.

NOTE: RUG-III is resource utilization groups, version III.

SOURCES: White, C., Harvard University, Pizer, S.D., Boston University and Department of Veterans Affairs, and White, A.J., Abt Associates, 2002.

diem staff-time costs are $104.66, and mean per diem non-therapy ancillary costs are $74.02. As expected, all three cost measures are right-skewed, with non-therapy ancillary costs being extremely skewed. The correlation coefficient for staff-time costs and non-therapy ancillary costs is 0.23.

As shown in Table 2, the RUG-III classification system explains only 10.4 percent of the variance in total costs (staff time plus non-therapy ancillary costs) among Medicare stays. Fries et al. (1994) reported that RUG-III explained 55.5 percent of the variance in costs. The gap between our results and these previously published results is primarily due to two differences in study design: we limit our sample for this analysis to Medicare stays, and we include non-therapy ancillary costs in our total cost variable. As shown in Table 2,

RUG-III performs much better when staff-time costs are used as the dependent variable instead of total costs, explaining 20.6 percent of the variance. Variance explanation increases again to 40.1 percent when staff-time costs are the dependent variable and all 3,793 stays (Medicare and non-Medicare) are included in the analytic sample. Some of the remaining gap in variance explanation (40.1 versus 55.5 percent) might be due to the fact that Fries et al. (1994) used self-reported staff time whereas our STM data were generated using electronic wands and bar codes. Also, if a case-mix classification system is developed using a particular data set, that classification system is likely to perform better on that data set than on an entirely new data set. This reflects the simple fact that a new data set will likely be drawn from a different population (different States or different

Table 3

Results of *t*-Tests Using Selected MDS-2.0 Items as Predictors of Residual Costs[1,2]

MDS-2.0 Item	Fraction of Residents with Item = "Yes" Percent	Mean Residual Costs Among Residents with Item = "No"	Mean Residual Costs Among Residents with Item = "Yes"
J1l: Shortness of Breath	15.6	-$11.76 (-$18.62 – -$4.90)	$63.79 ($37.32 – $90.27)
P1ag: Received Oxygen Therapy in Last 14 Days	19.0	-$15.97 (-$22.14 – -$9.81)	$68.02 ($41.89 – $94.15)
P1ac: Received IV Medication in Last 14 Days	14.1	-$11.95 (-$18.57 – -$5.33)	$72.74 ($42.79 – $102.70)
G6a: Bedfast All or Most of Time	13.7	-$12.12 (-$18.45 – -$5.79)	$76.18 ($43.38 – $108.97)
P1bda: Received Respiratory Therapy on 1 or More Days in Last 7 Days	12.4	-$17.23 (-$23.10 – -$11.35)	$121.45 ($85.29 – $157.62)
P1ai: Received Suctioning in Last 14 Days	3.9	-$5.59 (-$11.93 – -$0.76)	$137.25 ($43.24 – $231.25)
P1aj: Received Tracheostomy Care in Last 14 Days	2.6	-$5.21 (-$11.54 – $1.12)	$194.57 ($61.46 – $327.68)

[1] MDS-2.0 items are selected for presentation based on two criteria: a *p*-value <0.01 on the t-test of the difference in total costs between those with the item ="Yes" and those with the item = "No" (assuming unequal variance); and an absolute value of the difference between mean residual costs among residents with item = "Yes" and item = "No" greater than $75.

[2] Residual cost is the difference between an individual's cost and the mean costs within that individual's RUG-III category. Ninety five percent confidence intervals (shown in parentheses) are calculated using the standard error of the mean generated by the SAS/STAT *t*-test procedure.

NOTES: MDS-2.0 is minimum data set, version 2.0. N=1,304.

SOURCES: White, C., Harvard University, Pizer, S.D., Boston University and Department of Veterans Affairs, and White, A.J., Abt Associates, 2002.

years) and will likely be gathered and processed in a slightly different way (such as different outlier exclusion rules or different wage weights).

High-Cost Individuals within RUG-III Groups

Table 3 presents results of the *t*-tests using MDS-2.0 items as predictors of residual costs. Items were selected for presentation based on two criteria: a *p*-value <0.01, and an absolute difference greater than $75 in residual costs between residents with and without the MDS-2.0 item. This list of MDS-2.0 items is dominated by high-tech special treatments such as ventilator and respiratory therapy, suctioning, receiving intravenous medications, parenteral feeding, and tracheostomy care.

Many of the items listed are qualifying criteria for the extensive services major category. This led us to test whether, among residents assigned to the rehabilitation category, residents who also qualify for the extensive services category have

higher costs than those who do not. As shown in Table 4, those rehabilitation residents who also qualify for extensive services have somewhat higher staff-time costs and much higher non-therapy ancillary costs and total costs (*p*-values <0.001).

Non-Hierarchical Versions of RUG-III

As shown in Table 5, the non-hierarchical 44-dummy variable version of RUG-III explains much more of the variance in costs (24.9 percent) than the standard RUG-III (10.4 percent). The simple non-hierarchical version, with four dummy variables, also explains more of the variance in costs (21.1 percent) than RUG-III.

DISCUSSION

The main complaint from providers regarding the SNF PPS is that SNFs are undercompensated for dual needs residents who incur high rehabilitation costs and high non-therapy ancillary costs. This

Table 4
Mean Costs Among Residents Assigned a Rehabilitation Category

Item	Extensive Services Category	
	Does Not Qualify	Qualifies
Number of Residents	737	106
Staff Time Costs	$111.54	$139.04
	($108.40 - $114.68)	($127.59 - $150.49)
Non-Therapy Ancillary Costs	$50.57	$190.04
	($45.00 - $56.14)	($139.59 - $240.49)
Total Costs	$217.99	$384.97
	($211.16 - $224.82)	($330.58 - $439.35)

NOTES: Ninety five percent confidence intervals (shown in parentheses) are calculated using the standard error of the mean generated by the SAS/STAT t-test procedure. The t-tests generate p-values <0.001 on the difference between those who qualify for extensive services and those who do not for all cost measures.

SOURCES: White, C., Harvard University, Pizer, S.D., Boston University and Department of Veterans Affairs, and White, A.J., Abt Associates, 2002.

Table 5
Variance Explanation of Modified Versions of RUG-III

RUG-III Version	Variance Explanation[1]
	Percent
Standard	10.4
44-Variable Non-Hierarchical	24.9
Simple Non-Hierarchical	21.1

[1] Percent of variance explained is measured using the R^2 statistic from a regression with total costs as the dependent variable. The non-hierarchical modified versions of RUG-III are described in the text.

NOTES: N=1,304. RUG-III is resource utilization groups, version III. To create the 44-variable non-hierarchical model, the 44 mutually exclusive RUG-III groups are replaced by 44 dummy variables, each set to 1 or 0 depending on whether the resident met the qualification criteria for a particular RUG-III group. To create the simple non-hierarchical model, four dummy variables are created based on whether the resident met the qualification criteria for a group or groups in the following major categories: rehabilitation, extensive services, special care domain, and clinically complex.

SOURCES: White, C., Harvard University, Pizer, S.D., Boston University and Department of Veterans Affairs, and White, A.J., Abt Associates, 2002.

is not surprising, given that RUG-III was developed using a data set that only included staff time (Fries et al., 1994). The original developers of the RUG-III system did not measure non-therapy ancillary costs and assumed, out of some combination of necessity and convenience, that these costs were small and were closely associated with nursing costs (Fries et al., 1994). Non-therapy ancillary costs, in our sample of Medicare stays, are quite large (mean $74.02) and are poorly correlated with staff-time costs (correlation coefficient of 0.23). It is not surprising, therefore, that a classification system designed to predict staff time costs performs poorly when used to predict total costs.

The fact that residual costs are strongly associated with high-tech special treatments means that there are predictable cost patterns not accounted for in the RUG-III system. The most obvious pattern is that, among residents assigned to the rehabilitation category, no case-mix adjustment is made based on the special treatments (Table 3). This feature of the RUG-III system encourages SNFs to avoid certain types of residents, and also has implications for the distribution of payments across SNF types. A SNF with a high-severity resident population, such as a hospital-based SNF providing intravenous medications to many of its residents, is likely to fare poorly under the new PPS.

An argument against a non-hierarchical model, such as the 44-variable model we tested, is that there would be many more payment rates than there are in the current system, making the system more complex. One response to this argument is that a 44-variable non-hierarchical system can be thought of as having one base payment amount and 43 adjuster amounts—this is hardly more complex than the current system. The second, more weighty response is that a simple 4-variable non-hierarchical

model, with only 16 (2 to the fourth power) payment rates, explains more variance in costs than the 44-category RUG-III. The benefit of the non-hierarchical models comes, not from a profusion of rates or using variables that are excluded from RUG-III or from having a complicated structure, but from taking into account both rehabilitation therapy and special treatments received.

CMS, in response to this research and other evidence that RUG-III does poorly at identifying those residents with high non-therapy ancillary costs, (U.S. General Accounting Office, 1999) is taking steps to replace the RUG-III case-mix classification system. The modified versions of RUG-III tested in this article are not meant as proposed replacements; they simply highlight the weaknesses of RUG-III and show that increased variance explanation need not require increased complexity. The development of a new case-mix classification system would require more-recent cost data and a large sample of residents that represents the entire SNF population. As noted by CMS, there are serious technical challenges in developing and validating a new case-mix classification system (*Federal Register*, 2000a). These challenges are compounded by the fact that the Balanced Budget Refinement Act of 1999 and the Benefits Improvement and Protection Act of 2000 introduced temporary payment increases that will be rescinded as soon as CMS implements a new case-mix classification system (*Federal Register*, 2000b; 2001). The SNF industry is, therefore, reluctant to endorse any change. Nevertheless, the ongoing CMS-funded research is likely to result in the development and implementation of a significantly improved case-mix classification system for SNFs.

REFERENCES

Employee Compensation, Table 15. Private Industry, Health Services by Industry and Occupational Group. Bureau of Labor Statistics, 1999. Internet address: http://www.bls.gov/news.release/ecec.t15.htm.

Federal Register: Medicare Program; Prospective Payment System and Consolidated Billing for Skilled Nursing Facilities; Final Rule. 63FR(91):26098-26147 (May 12, 1998).

Federal Register: Medicare Program; Prospective Payment System and Consolidated Billing for Skilled Nursing Facilities–Update; Final Rule. 65 FR (147):46769-46796 (July 31, 2000a).

Federal Register: Medicare Program; Prospective Payment System and Consolidated Billing for Skilled Nursing Facilities-Update; Proposed Rule. 65 FR (69):19187-19236 (April 10, 2000b).

Federal Register: Medicare Program; Prospective Payment System and Consolidated Billing for Skilled Nursing Facilities–Update; Proposed Rule. 66 FR (91):23983-24036 (May 10, 2001).

Fries, B.E., Schneider, D.P., Foley, W.J., et al.: Refining a Case-Mix Measure for Nursing Homes: Resource Utilization Groups (RUG-III). *Medical Care* 32(7):668-685, 1994.

Health Care Financing Administration: *Long Term Care Facility Resident Assessment Instrument (RAI) User's Manual [Official Replica Edition]*. Natick, MA. Eliot Press, 1999.

Health Care Financing Administration: *SAS Code for RUG3 Version 5.12 (44 Group Model), 2000*. Internet address: http://cms.hhs.gov/medicaid /mds20/ mdssoftw.asp#512

Newhouse, J.P.: Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection. *Journal of Economic Literature* 34:(3) 1236-1263, September 1996.

U.S. Congress: *Balanced Budget Act of 1997*, 105th Congress, 1997. Public Law 105-33, enacted August 5. U. S. Government Printing Office. Washington, DC. 1997.

U.S. Department of Labor: *Employer Costs for Employee Compensation, Table 15, Private Industry, Health Services by Industry and Occupational Group*. Bureau of Labor Statistics, 1999. Internet address: http://www.bls.gov/news.release/ecec.t15.htm

U.S. General Accounting Office: *Skilled Nursing Facilities: Medicare Payments Need to Better Account for Nontherapy Ancillary Cost Variation*. GAO/HEHS-99-18. U.S. Government Printing Office. Washington, DC, 1999.

U.S. General Accounting Office: *Skilled Nursing Facilities: Providers Have Responded to Medicare Payment System by Changing Practices*. GAO-02-841. U.S. Government Printing Office. Washington, DC. August 2002.

U.S. Government Printing Office: *Public Law 105-33*. U. S. Government Printing Office. Washington, DC. 1999.

White, A., Pizer, S., and White, C.: *Refining Resource Utilization Groups (RUG-III) for a National Skilled Nursing Facility System, Technical Expert Panel Briefing Book*. Abt Associates. Cambridge, MA, 1999.