Final Report to CMS

Options for Improving Medicare Payment for Skilled Nursing Facilities

Centers for Medicare and Medicaid Services
CMS Contract No. 500-00-0025
Project Officer: Jeanette Kranacs

First submitted: August 2004
This version: March 2007

Project Team
Principal Investigator: Korbin Liu

Urban Institute
Bowen Garrett
Korbin Liu
Sharon Long
Stephanie Maxwell
Yu-Chu Shen
Douglas Wissoker

University of Colorado Health Sciences Center
Theresa Eilertsen
Anne Epstein
Andrew Kramer (lead investigator, UCHSC)
Sung-Joon Min
Robert Schlenker

University of Michigan
Brant Fries

Harvard University
Joan Buchanan

Research in this report was supported by the Center for Medicare and Medicaid Services (CMS) Contract No. 500-00-0025 (UI #07108-002-00) "Assessment, Refinement and Analysis of the Existing Prospective Payment System for Skilled Nursing Facilities." Inferences and opinions in this report are solely those of the authors and do not necessarily represent the views of any of the institutions with which they are affiliated, or CMS.

Table of Contents

I. Introduction 1
II. Background 5
A. Reimbursement Policies.................... 5
B. RUG-III Classification System 7
C. Areas for Potential Improvement in the Existing SNF PPS.................... 9
Non-Therapy Ancillaries 9
Therapy Payments Based on Care Used 10
Developing PPS Payment Weights: The Staff-Time Measurement Method.................... 11
III. Data and Methods 14
A. Data Sources 14
B. Samples and "Base Case" Facility and Stay Data Exclusions 16
Application of cost to charge ratios 17
C. Analysis.................... 17
D. Explanatory Variables.................... 18
E. Analytical Issues.................... 20
Coding of Diagnoses.................... 20
Ambiguity in the Meaning of MDS Service Use.................... 21
The Small Size of the Medicare Share in Most SNFs 21
F. Statistical Evaluation of Models.................... 22
Patient Stay-Level Models and Statistics.................... 22
CART (Classification and Regression Tree Analysis) 23
R-squared.................... 24
Sensitivity 24
Standard Deviation of the Relative Weights.................... 25
Validation Out-of-Sample.................... 25
Facility-Level Models.................... 26
Facility Case-Mix Indices.................... 27
Facility-Level Model R-squared.................... 28
The CMI Coefficient and CMI Compression 28
Interpreting Compression Results for the Payment and Fully-Specified Models.................... 30
IV. RUG-Based Approaches to Classifying NTA Costs 31
A. Aims and Background.................... 31
B. Data and Methods.................... 33
Exploratory Findings 33
Models of NTA Costs and Charges 36
C. Findings.................... 37
Bivariate Relationships Between Model Variables and NTA Costs 38
Multivariate Models.................... 41
Highlighting Variance Explanation of Different Models 42
Application of SIM Models to Create Payment Cells 45
D. Implications at the Facility Level.................... 46
Models with Cost-Based Weights and Case-Mix Index.................... 47
Person-Level Models 49

Facility-Level Models 49
Models with Charge-based Weights and Case-mix Index 51
Person-Level Models 53
Facility-Level Models 53
E. Additional Findings 54
F. Discussion 55
Statistical Significance 57
Clinical Meaningfulness 57
Administrative Burden 57
Clinical and Financial Incentives 58
Ease of Implementation and Administration 58
V. New Profiles Approach 59
A. Aims and Background 59
B. Development of the New Patient Categories 60
Conceptual Development 60
Operationalizing the Group Definitions 62
Discussion 70
C. NP-NTA Models 71
Data 73
NTA Resource Use Variables 73
Explanatory Variables 76
NP patient classification system 76
Other explanatory variables 77
Methods 78
Two-Stage Regression Models 78
Regression Tree (CART) Models 79
Results of the Initial Models 80
Model Fit 80
Variables Associated with NTA Charges in the Two-Stage Models 82
Variables Related to Resource Use in CART Models 86
Moving the Models toward a Payment System 86
Inappropriate Payment System Variables 87
Consolidating the NTA Components into a Single Model 88
Refining the Variable Selection Process 89
Final Models 90
Validation and Facility-Level Analyses 95
Facility Case Mix Index (CMI) Models 98
D. Discussion 100
E. NP-Therapy Models 101
Methods 103
F. Results 105
Distributions 105
Preliminary Models 111
Final Therapy Cost Models 118
Validation and Facility-Level Analyses 125
Discussion 130

VI. DRG-Based Approaches to Classifying Medicare SNF Patients 135
A. Aims and Background 135
B. Data and Methods 139
DRG of Qualifying Hospital Stay of SNF Patients 140
Creating a Functioning Model for Use with DRGs 141
Stay-Level Regression Models 143
Facility-Level Regression Models 144
C. Findings 145
Summary of Cost and Charge Measures 145
Description of DRGs in SNFs 147
Description of Functional Variables and Relationship with SNF Costs 150
Predictive Ability of DRGs and Functioning Variables for SNF Costs at the Stay-Level 152
Predictive Ability of DRGs and Functioning Variables for SNF Charges at the Stay-Level ... 155
Correlations between Cost- and Charge-Based Relative Weights 157
Facility-Level SNF Cost Models Using Cost-Based CMIs 158
Facility-Level SNF Cost Models Using Charge-Based CMIs 161
D. Discussion 163
Statistical Significance 164
Clinical Meaningfulness 165
Administrative Burden 165
Financial Incentives 166
Conclusion 167
VII. Comparison of Models 169
A. Introduction 169
B. NTA-Specific Options 170
Models 171
RUG-58 171
Service Index Model (SIM) 172
New Profile NTA Model (NP-NTA) 173
Statistical Comparison of the three NTA Models 175
Comparison of Person-Level NTA Models Based on Costs 175
Comparison of Person-Level NTA Models Based on Charges 177
Comparison of Facility-Level Models Based on Cost Weights 178
Comparisons of Facility-Level NTA Models Based on Charge Weights 181
Summary of Statistical Comparisons 183
Comparisons of NTA Models Based on Other Policy Criteria 184
Clinical Meaningfulness 184
Ease of Implementation and Administration 185
Ease of Use by Providers 185
Appropriate Clinical and Financial Incentives 186
C. Rehabilitation Therapy Specific Options 186
New Profile Rehabilitation Therapy Model (NP-Therapy) 186
Statistical Comparison of “Need Based” and RUG-III Rehab Therapy Options 188
Comparison of Person-Level Rehabilitation Cost Models 188
Comparison of Facility-Level Rehab Cost Models 190
Comparison of NP-Therapy Models and RUG-III Based on other Policy Criteria 192

Clinical Meaningfulness ... 192
Ease of Implementation and Administration ... 192
Ease of Use by Providers ... 193
Appropriate Clinical and Financial Incentives ... 193
D. DRG-Related Options ... 193
Model ... 194
Statistical Comparison of DRG-Based and RUG-III Models ... 195
Comparison of Person-Level Cost Models ... 195
Comparison of Facility-Level Cost Models ... 198
Comparison of DRG-Based Models and RUG-III Models Based on Other Policy Criteria ... 201
Clinical Meaningfulness ... 201
Ease of Implementation and Administration ... 201
Ease of Use by Providers ... 202
Appropriate Clinical and Financial Incentives ... 202
E. Discussion ... 202
VIII. SNF PPS Outlier Payment Policies ... 204
A. Introduction ... 204
B. Components of Medicare PPS Outlier Payment Policies ... 207
Outlier Target ... 207
Loss Amount ... 208
Loss-Sharing Ratio ... 209
C. Data Sources ... 211
D. Outlier Payment Policy Approaches and Parameters Analyzed for the SNF PPS ... 212
Budget-Neutral Policies ... 212
Total Cost Outlier Policies ... 212
NTA Cost Outlier Policies ... 213
Loss-Share Approaches ... 215
Using per Stay Versus per Diem Costs in Determining Outlier Payments ... 216
E. Total Cost Outlier Policy Findings ... 217
Summary Statistics ... 218
Outlier Stay and Payment Percentages, by RUG-III Category ... 221
Outlier Stay and Payment Percentages, by Facility Characteristic ... 222
Outlier Payments per Stay, by Facility Characteristic ... 228
Ratio of Total Payments to Costs, by Facility Characteristic ... 232
F. NTA Cost Outlier Policy Findings ... 235
Summary Statistics ... 236
Outlier Stay and Payment Percentages, by RUG-III Category ... 239
Outlier Stay and Payment Percentages, by Facility Characteristic ... 241
Outlier Payments per Stay, by Facility Characteristic ... 245
Ratio of Total Payments to Costs, by Facility Characteristic ... 249
G. Additional Comparisons between Total and NTA Policy Approaches ... 251
Overlap of Stays Receiving Outlier Compensation under Total and NTA Policies ... 251
Comparing Total and NTA Policies that Each Compensate 10% of SNF Stays ... 253
H. Concluding Comments ... 257
IX. Implications for Research and Policy ... 260

A. Research 260
MDS Variables on Service Use 260
Hospital and SNF Diagnoses on Claims 261
Disconnect between MDS Assessment and Claims Data 262
Hospital-Based and Freestanding SNFs 263
B. Policy 263
How Much Change? 265
Which Components? 265
Choosing Among Different Models 267
References 269
Appendix to Chapter II 273
Appendix II.1: Variance Explanation of RUG-III 273
Appendix to Chapter III 274
Appendix III.1: Data and Methods 274
Data Sources 274
Construction of the 2001 Analysis Files 274
Facility Data Exclusion Rules (Flags) 276
Unreliable Cost Data 276
Inability to Assign Medicare Costs in a Facility 276
Unreliable Cost to Charge Ratios 276
Stay Data Exclusion Rules (Flags) 277
Selecting Units of Observation 277
Appendix III.2: Statistical Evaluation of Patient Classification Models 279
Patient Stay-Level Models and Statistics 279
R-squared 280
Sensitivity 280
Standard Deviation of the Relative Weights 280
Validation Out-of-Sample 281
Facility-Level Models 281
Facility Case-Mix Indices 282
Facility-Level Model R-squared 283
The CMI Coefficient and CMI Compression 283
Interpreting Compression Results for the Payment and Fully-Specified Models 286
Appendix to Chapter VIII 290

List of Tables and Figures

Table III.1 Number of Cases Remaining after Specific Initial Edits 16
Table III.2: Number of SNF Stay Observations in the 10% Stay and 10% Facility/Stay Analysis Samples 17
Table III.3: Average Characteristics of SNF Patients, 2001 19
Table IV.1: Average Wage-Adjusted NTA Stay Cost Per Diem by Presence of MDS and Claim Report, by Assessment Number 35
Table IV.2: Coefficients from Bivariate Regressions of NTA Cost Per Diem on WIM Variables, Estimated Using 2001 DataPRO Test File 40
Table IV.3: Coefficients from Bivariate Regressions of NTA Cost Per Diem on SIM Variables, Estimated Using 2001 DataPRO Test File 40
Table IV.4: SIM Model for Predicting Adjusted NTA Costs Per Diem – Estimated Using 2001 DataPRO Test File 41
Table IV.5: Comparison of Predictive Models Incorporating RUG-III, RUG-58, WIM and SIM, Estimated Using 2001 DataPRO Test File 43
Table IV.6: Groupings and Payments from CART Model Estimated Using 2001 DataPRO Test File 46
Table IV.7: Comparison of Predictive Models of NTA Costs: RUG-III, RUG-58, and SIM+RUG58G[a] 48
Table IV.8: Comparison of Predictive Models of NTA Charges: RUG-III, RUG-58, SIM+RUG58G[a] 52
Figure V.1: Relationship between Function and Therapy 64
Table V.1: Potential Acute Condition ICD-9-CM Codes 65
Figure V.2: New SNF Patient Classification System 67
Table V.2: Selected Characteristics of the Three Patient Groups in a New SNF Patient Classification System 69
Table V.3: Average Per Diem Charges and Wage-Adjusted Costs of the Three Patient Groups in a New SNF Patient Classification System 70
Figure V.3: Drug Cost Distribution (All Stays with Costs > $0) 74
Figure V.4: Respiratory Cost Distribution (All Stays with Costs > $0) 74
Figure V.5: ONTA Cost Distribution (All Stays with Costs > $0) 75
Table V.5: Explanatory Power for All Models 82
Table V.6: Variables Associated with at Least $20 of Drug Charges Per Day 83
Table V.7: Variables Associated with at Least $15 of Respiratory Charges Per Day 84
Table V.8: Variables Associated with at Least $15 of ONTA Charges Per Day 85
Table V.9: Variables Excluded from the Model for Payment Related Reasons 88
Table V.10: Means for Variables in the Final NTA Models 91
Table V.11: Final NTA Models: Variables and Their Dollar Effects on Cost Per Day 92
Table V.12: Performance Summary for NP NTA Stay-Level Models 95
Table V.13: Performance Summary for Stay- and Facility-Level NTA Cost Models 96
Table V.14: Performance Summary for Stay- and Facility-Level NTA Charge Models 97
Figure V.7: PT Charge Distribution (All Stays with PT Charges > $0) 106
Figure V.8: OT Charge Distribution (All Stays with OT Charges > $0) 107
Figure V.9: ST Charge Distribution (All Stays with ST Charges > $0) 108
Figure V.10: PT/OT Charge Distribution (All Stays with PT/OT Charges > $0) 109

Table V.18: Distribution Skew for Therapy Charges 110
Table V.19: Variation in Therapy Cost and Charges Per Day.. 111
Table V.20: Correlations of Predictors with PT/OT vs. ST Charges.. 113
Table V.21: Explanatory Power for Initial Charge Models 114
Table V.22: Variables Associated with at least $5 of PT/OT Charges per Day 116
Table V.23: Variables Associated with at least $1 of ST Charges per Day 118
Charges R-squared 118
Table V.24: Means for Variables in the Final PT/OT and ST Cost Models............................. 119
Table V.25: Final PT/OT Models: Variables and their Dollar Effects on Cost per Day.......... 122
Table V.26: Final ST Models: Variables and Their Dollar Effects on Cost per Day 124
Table V.27: Performance Summary for Stay and Facility-Level PT/OT Cost Models............. 126
Table V.28: Performance Summary for Stay and Facility-Level ST Cost Models 129
Figure VI.1: Total SNF Cost Per Day Relative Weight vs. Acute Hospital Relative Weight (Top 30 DRGs in SNFs 1999) 137
Table VI.1: Summary of Wage-Adjusted SNF Cost and Charge Measures............................. 146
Table VI.2: 30 Most Common DRGs in SNFs 148
Table VI.3: 30 Most Costly DRGs in SNFs (With at Least 100 Cases in the Test Sample)..... 149
Table VI.4: Description of Functional Variables.. 151
Table VI.5: Cost Per Day Regression Results for Functional Variables 151
Table VI.6: Models of Person-Level SNF Costs by Component Using DRGs + Functioning as Explanatory Variables.. 153
Table VI.7: Models of Person-Level SNF Charges by Component Using DRGs + Functioning as Explanatory Variables.. 156
Table VI.8: Correlations Between Cost and Charge Relative Weights 157
Table VI.9: Models of Facility-Level Costs Using Cost-Based DRGs + Functioning CMIs[a] .. 159
Table VI.10: Models of Facility-Level Costs Using Charge-Based DRGs + Functioning CMIs[a] 162
Table VII.1: Comparison of Person-Level NTA Costs Models: RUG-III, RUG-58, SIM+RUG-58G[a], and NP-NTA 176
Table VII.2: Comparison of Person-Level NTA Charge Models: RUG-III, RUG-58, SIM+RUG-58G[a], and NP-NTA 177
Table VII.3: Comparison of Facility-Level NTA Costs Models: RUG-III, RUG-58, SIM+RUG58G[a], and NP-NTA 179
Table VII.4: Comparison of Facility-Level NTA Charges Models: RUG-III, RUG-58, SIM+RUG-58G[a], and NP-NTA 182
Table VII.5: Comparison of Person-Level Rehabilitation Therapy Costs Models: RUG-III and NP-Therapy 189
Table VII.6: Comparison of Facility-Level Rehabilitation Therapy Costs Models: RUG-III and NP-Therapy 190
Table VII.7: Comparison of Person-Level SNF Total Costs Models, RUG-III and DRG + Functioning 197
Table VII.8: Comparison of Facility-Level Total Costs Models: RUG-III and DRG + Functioning 199
Table VIII.1: Medicare Prospective Payment Systems: Outlier Policy Parameters, 2004....... 208
Table VIII.2: Summary Statistics under Five Total Outlier Payment Policy Simulations 218

Table VIII.3: Summary Statistics under Five Total Outlier Payment Policy Simulations, (Using Alternative Loss-Share Methodology) 220
Table VIII.4: Outlier Stay and Payment Percentages under Five Total Outlier Payment Policies, by RUG-III Category 221
Table VIII.5: Outlier Stay and Payment Percentages under Five Total Outlier Payment Policies, by Facility Characteristics 224
Table VIII.6: Outlier Stay and Payment Percentages under Five Total Outlier Payment Policies, by Facility Characteristics (Using the "Separate Median" Loss-Share Methodology) 226
Table VIII.7: Average Outlier Payments Across all Stays under Five Total Outlier Payment Policies, by Facility Characteristics 229
Table VIII.8: Average Outlier Payments Across all Stays under Five Total Outlier Payment Policies, by Facility Characteristics (Using the "Separate Median" Loss-Share Methodology) 231
Table VIII.9: Ratios of Total Payments to Costs under Five Total Outlier Payment Policies by Facility Characteristics 233
Table VIII.10: Summary Statistics under Eight NTA Outlier Payment Policy Simulations 237
Table VIII.11: Summary Statistics under Eight NTA Outlier Payment Policy Simulations (Using the "Separate Median" Loss-Share Methodology) 239
Table VIII.12: Outlier Stay and Payment Percentages under Three NTA Outlier Payment Policies, by RUG-III Category 240
Table VIII.15: Average Outlier Payments Across all Stays under Three NTA Outlier Payment Policies, by Facility Characteristics 246
Table VIII.16: Average Outlier Payments Across all Stays under Three NTA Outlier Payment Policies, by Facility Characteristics (Using the "Separate Median" Loss-Share Methodology) 248
Table VIII.17: Ratios of Total Payments to Costs under Three NTA Outlier Payment Policies, by Facility Characteristics 250
Table VIII.18: Percent of Stays Receiving Outlier Payments under Three NTA Policies that also Receive Outlier Payments under Total Policies 252
Table VIII.19: Average Costs and Payments per Stay of Stays Receiving Outlier Payments under the 4% Total and 15% NTA Outlier Policies 254
Table VIII.20: RUG-III Category of Stays Receiving Outlier Payments under a 4% Total Outlier Policy and a 15% NTA Outlier Policy 255
Table VIII.21: Facility Characteristics of Stays Receiving Outlier Payments under the 4% Total and 15% NTA Outlier Policies 256

List of Appendix Exhibits

Appendix Exhibit III.1: The Number of Cases Remaining after Specific Initial Edits 275
Appendix Exhibit III.2: Illustration of CMI Coefficient and Facility CMI Scenarios 285
Appendix Exhibit III.3: CMI Coefficient Scenarios in Payment and Fully-Specified Facility Models and Implications 287
Appendix Exhibit VIII.1: Distribution of SNF Total Costs per Stay, 2001 290
Appendix Exhibit VIII.2: Distribution of SNF Total Costs Per Day, 2001.............................. 291
Appendix Exhibit VIII.3: Distribution of SNF NTA Costs Per Stay, 2001 292
Appendix Exhibit VIII.4: Distribution of SNF NTA Costs per Day, 2001 293
Appendix Exhibit VIII.5: List of States by Census Division.. 294
Appendix Exhibit VIII.6: Outlier Stay and Payment Percentages under Five Total Outlier Payment Policies, by RUG-III Group 295
Appendix Exhibit VIII.7: Outlier Stay and Payment Percentages under Eight NTA Outlier Payment Policies, by RUG-III Category 296
Appendix Exhibit VIII.8: Outlier Stay and Payment Percentages under Eight NTA Outlier Payment Policies, by RUG-III Group 297
Appendix Exhibit VIII.9: Outlier Stay and Payment Percentages under Eight NTA Outlier Payment Policies, by Facility Characteristics 298
Appendix Exhibit VIII.10 299
Appendix Exhibit VIII.11: Average Outlier Payments Across all Stays under Eight NTA Outlier Payment Policies, by Facility Characteristics 300
Appendix Exhibit VIII.12: Average Outlier Payments Across all Stays under Eight NTA Outlier Payment Policies, by Facility Characteristics (Using Alternative Loss-Share Methodology) 301
Appendix Exhibit VIII.13: Ratios of Total Payments to Costs under Eight NTA Outlier Payment Policies, by Facility Characteristics 302

I. Introduction

After nearly a decade of double-digit rates of growth in Medicare expenditures for skilled nursing facility (SNF) services, the Balanced Budget Act (BBA) of 1997 mandated that the program's cost-based, retrospective reimbursement policy for SNFs be replaced by a prospective payment system (PPS). Initiated in 1998, the Medicare SNF PPS established a prospectively determined per-diem payment rate for SNF patient care— adjusted for case-mix, area wages, urban or rural status, and changes in input prices.[1]

The SNF PPS has generally accomplished its major objective of curbing Medicare spending growth for SNF services. The new system provides a full range of SNF services to Medicare beneficiaries while allowing providers more flexibility in their use of Medicare funds. In addition, studies of the impact of the SNF PPS do not currently indicate systematic problems with beneficiary access or quality of care (Maxwell, et al. 2003).

Despite its accomplishments, stakeholders and policy analysts express varying degrees of concern over three aspects of the existing SNF PPS: (1) the ability of the current patient classification system to adequately account for cost variations of non-therapy ancillary (NTA) services, such as prescription medicines; (2) the basis of payment for rehabilitation therapy services on the amount of services used, rather than on expected need as reflected by patient characteristics; and (3) the ability of the relative payment weights for nursing services to reflect current care practices.

---

[1] A market basket index is applied which accounts for changes over time in the prices of an appropriate mix of goods and services included in covered SNF services.

Seeking possible improvements in the SNF PPS, Congress passed a provision in the 2000 Benefits Improvement and Protection Act (BIPA) requiring that the Centers for Medicare and Medicaid Services (CMS) conduct research to assess potential refinements and alternatives to the existing payment system. In response to the Congressional mandate, CMS is funding a multi-year study conducted by analysts at the Urban Institute, University of Colorado, University of Michigan, and Harvard University.

In the first three years of the study, we conducted extensive research to identify determinants of Medicare SNF costs. We analyzed data from multiple sources, including SNF and prior hospital claims, minimum data set (MDS) assessments, and SNF cost reports. Our findings fostered development of multiple patient classification models for different cost components, including NTA services and rehabilitation therapy services.

Our initial research findings were discussed previously in two interim reports. The first report published in 2003 described the development of the current PPS, analyzed the main issues with the SNF PPS, and assessed the literature to date regarding the SNF PPS' impact on patient access to SNF services, patient outcomes, and providers' financial performance under the new payment system (Maxwell, et al. 2003). The second report published in 2004 identified determinants of Medicare SNF costs and developed person-level classification models of SNF patients' costs (Urban Institute, 2004). In addition, several Technical Advisory Panel (TAP) meetings were held in order to obtain input and suggestions for further study.

This report offers potential public policy options to refine Medicare's payment of SNF services, by developing five specific patient classification models. Three of the models classify patients according to NTA costs. The first two of these are extensions of prior research supported by CMS (Abt Associates, 2000), using data only from SNF claims and MDS to model total NTA costs. We refer to these as RUG-based models. The third model takes a new approach to classifying patients according to components of their NTA costs (i.e., drugs, respiratory services, other), using data from prior hospital stays as well as SNF claims and the MDS. We refer to this as New Profile NTA (NP-NTA). The fourth model, also totally new, was developed to predict patients' need for rehabilitation therapy services, so payment for therapy services could be based on patient characteristics than on the actual use of therapy services. We refer to this as New Profile Rehabilitation Therapy (NP-Therapy). The fifth is a diagnostic related group (DRG)-based model addressing total (including nursing) costs of SNF care. Associated DRG-based models for NTA and rehabilitation therapy components were also developed under this classification strategy.

We also developed various outlier payment options to address unforeseen or extraordinarily costly cases. Such policies can be applied to the existing SNF PPS or in conjunction with any of the five patient classification models that we developed.

It is possible to "mix and match" the different classification models with the outlier policy choices in order to produce further options for improving the SNF PPS. For example, any of the NTA models could independently address the perceived problems of inadequate accounting for the distribution of NTA costs; yet if paired with the NP-Therapy mode, the newly acquired model could address the rehabilitation therapy fee schedule problem as well.

The analysis we present offers an empirical investigation of the SNF PPS system currently in place, highlighting potential areas of improvement at the policy level. In brief, Chapter II of this report provides background on the Medicare SNF provider reimbursement policies, and three key areas for potential improvement within the current SNF PPS. Chapter III briefly describes the data sources and methodological issues. A more thorough explanation of the data and methodology used, in addition to figures and tables supporting the claims herein, can be found in the Appendix to this report. Chapters IV through VI discuss the five patient classification models that potentially constitute the "building blocks" for improving the SNF PPS. Chapter VII compares the pros and cons of the five models. Chapter VIII discusses and simulates options for implementing an outlier payment policy. Chapter IX concludes the report with a discussion of the research and policy implications of our work.

II. Background

In this chapter, we first discuss Medicare's SNF payment policies and spending trends leading to the Balanced Budget Act of 1997. We then describe the RUG-III classification system, the basis for adjusting for case-mix in the SNF PPS that was implemented in response to the BBA mandate. Finally, we review three areas for potential improvement in the SNF PPS, as noted by some policy analysts, providers, and consumer groups.

Until the mid 1980's, SNF services accounted for only a small percentage of total Medicare expenditures and were generally viewed as cost-effective and less intensive alternatives to extended acute-care hospital stays. After implementation of Medicare's acute-care hospital PPS in 1984, however, Medicare expenditures for SNF services began to grow rapidly. Between 1990 and 1996, for example, Medicare payments for SNFs rose from $2.5 billion to $11.7 billion. The rapid increase catalyzed concern among policy makers that use of these services had become excessive and did not necessarily improve the health of Medicare beneficiaries. Acting on these concerns, Congress enacted provisions in the 1997 BBA mandating that, in the future, Medicare SNF services should be paid under a PPS.

A. Reimbursement Policies

Prior to the BBA, Medicare reimbursed SNFs on three different bases, depending on three components of costs. In general, routine operating services were paid on an actual cost basis up to a per diem limit, ancillary services were paid on a reasonable cost basis, and capital was paid on a pass-through basis. Separate cost limits applied to hospital-based and freestanding

SNFs, and between urban and rural SNFs.[2] New providers were exempt from these limits for the first three years of operation. In addition, facilities could receive exception payments if they could demonstrate that their Medicare patient case-mix was sufficiently higher than average to warrant higher payments.

In the early 1990's, therapy and non-therapy ancillary service costs constituted a growing share of SNF expenditures and amounted to about half of all SNF payments by 1995. This trend was certainly not surprising, since ancillary services were not subject to per-diem cost limits. Most ancillary services were reimbursed under Medicare Part A. If patients were not covered by Part A or if they were not directly furnished by the SNF, some ancillary services could be reimbursed under Medicare Part B. A study conducted by The Healthcare Financing Administration estimated that in 1992, approximately 15% of therapy charges provided to SNF patients were billed to Part B (Liu 1993). In general, discontinuities in Part A and Part B accounting systems meant that Medicare could not readily monitor total program spending for SNF patients.

The 1997 BBA mandate moved SNFs into a per diem PPS that covered routine, ancillary, and capital costs—including items and services for which payments had previously been made under Part B, with a few exceptions (e.g., physician and psychologist services). From the 3-year transition period (1998-2001)[3] to the full implementation of the PPS, SNFs received payment

[2] Low volume SNFs (i.e., those with less than 1,500 Medicare days in a year) could elect to be paid at a rate equal to the lesser of the relevant limit or 105 percent of mean operating and capital costs of all (both hospital-based and freestanding) facilities in their region. This option was implemented to lessen the administrative burden on low use SNFs, with the goal of increasing access to SNF care for Medicare beneficiaries.

[3] The transition applied to facilities with cost reporting periods beginning on or after July 1, 1998 through cost reporting periods beginning on or after July 1, 2000.

derived concomitantly from: (a) a case-mix-adjusted Federal rate; and (b) a facility-specific rate based on the facility's historical costs. This blend steadily changed over the three year period in a way that weighted the Federal rate more heavily; payments reflected 25% of the Federal rate in the first year, 50% in the second year, 75% in the third year, and 100% thereafter.

The Federal rate, which became the full Medicare payment rate after 2001, was set at a level equal to a weighted mean of freestanding facility costs plus 50% of the difference between the freestanding mean and a weighted mean of all SNFs' (hospital-based and freestanding) costs. Separate rates were derived for SNFs in urban and rural areas, and further adjustments were made for case-mix and geographic variations in wage rates. The Federal rates were also adjusted to account for a facility's case-mix using the resident classification system, RUG-III. Subsequently, exception payments for case-mix were eliminated from the Federal rate.

B. RUG-III Classification System

The Resource Utilization-III model (RUG-III) was developed by Fries and colleagues (1994) based on a sample of 7,658 Medicare, Medicaid, and private-pay patients in 202 facilities across seven states. RUG-III is a hierarchical classification system with a structure that has forty-four groups falling into seven major categories:

- rehabilitation (14 final RUG-III groups)
- extensive services (3 groups)
- special care (3 groups)
- clinically complex (6 groups)

- impaired cognition (4 groups)
- behavior problems only (4 groups)
- reduced physical function (10 groups)

Virtually all Medicare SNF patients fall into the first 4 categories. In 2001, almost 79% of patient days were in the *rehabilitation* category, over 7% were in *extensive services*, and about 6% were in *special services* and *clinically complex* categories, respectively. The remaining three categories describe patients with primarily physical or cognitive disabilities, who generally do not constitute a significant part of the Medicare SNF population.

During the RUG-III development process, patients in the *rehabilitation* category were found to be the most resource intensive or costly; thus, that category became the foremost or highest category in the RUG-III hierarchy. This category is described mainly by the amount and use of physical, speech, and occupational therapy, and is further split into five sub-categories, which range from requiring a high of 720 total therapy minutes per week to a low of 45 minutes per week. The *extensive services, special care*, and *clinically complex* categories include patients who require NTA services such as IV medications, respiratory therapy, skin ulcer or surgical wound care, or care for other conditions such as pneumonia or dehydration. Patients in any of these categories may also concomitantly receive therapy , but at a low enough level that would not qualify them for one of the *rehabilitation* categories. Across most RUG-III categories, patients are divided into a final classification group based on their performance on an index of four Activities of Daily Living (ADLs): eating, toileting, bed mobility, and transferring.

C. Areas for Potential Improvement in the Existing SNF PPS

With varying degrees of consensus by issue, stakeholders and policy analysts, as noted, have identified three areas for potential refinement of the current SNF PPS. First, RUG-III does not adequately adjust for variations in costs of services, such as prescription medicines. Second, payments for rehabilitation therapy services (physical therapy, occupational therapy, and speech-language pathology services) are contingent upon the amount of service rendered, rather than on expected need based on patient characteristics. Third, the payment weights utilized for nursing services are based on special studies that may not reflect current care provisions. We briefly discuss these three issues in turn.

Non-Therapy Ancillaries

When the RUG-III system was first developed, NTA costs were not a large component of the costs of care in SNFs. But as the costs of prescription drugs and other NTA services have grown, they have become a larger component of SNF costs in recent years.

To this end, a potential improvement in the SNF payment system lies in modifying the current case-mix classification system to better reflect variations in NTA costs. At present, NTA payments in the SNF payment system are allocated across RUG-III groups in the same relative manner as that for nursing payments. While NTA resource use is somewhat correlated with the RUG-III nursing weights, it now varies dramatically more than nursing resource use—18-fold for SNF patients' NTA costs per day versus 2-fold for the RUG-III nursing component payment weights. As a result, the RUG-III system explains only about 5% of the variation in NTA costs among SNF patients (Kramer, et al. 1999, Abt Associates 2000, Liu, et al. 2002, Fries 2002).[4]

Therapy Payments Based on Care Used

Payment for rehabilitation therapy services under the current SNF PPS is based on the number of minutes of therapy, by therapy discipline, ordered by providers. This amount depends, in turn, on provider expectations, rather than on established relationships between patients' clinical and functional characteristics and the amount of therapy services needed. Accordingly, the therapy component of RUG-III functions more like a fee schedule than a traditional PPS. The implementation of a "fee schedule-like" mechanism for rehabilitation therapy evolved, in part, because of a desire to ensure adequate provision of therapy service to all recipients. In addition, prior research of nursing home patients in general suggested difficulty in finding health status or functional status characteristics that adequately explain therapy furnished to patients in nursing facilities.

This fee-schedule aspect of the SNF PPS has advantages and disadvantages. First, providers like the certainty of this aspect of the SNF PPS. In addition, the nature of this payment component helps ensure that services were actually provided when payments were made for rehabilitation therapy. Finally, because this system controls unit price rather than utilization of rehabilitation therapy services, it is viewed positively by some observers concerned about the general shortage of such services in nursing facilities.

---

[4] A more detailed review of variance explanation studies of the RUG-III system is presented in Appendix II-1 of the report and Maxwell, et al., 2003.

However, provider discretion to classify patients in a given rehabilitation group enables them to respond rapidly to changing payment policies in ways that appear related more to financial considerations than patient needs. The total payment rates for the rehabilitation groups are generally the highest of the RUG-III groups, creating strong incentives to assign patients to the rehabilitation category. For example, most patients are in the “medium” and “high” rehabilitation groups in RUG-III, which are the most profitable ones on a payment-per-minute of therapy basis. Studies (GAO 2002a) have shown that Medicare patient placement rates in these groups increased after the temporary payment add-ons were legislated by Congress in the years after the BBA.

Developing PPS Payment Weights: The Staff-Time Measurement Method

A third area for potential improvement in the SNF PPS relates to its method of deriving payment weights, which are based on a modern variant of industrial time and motion studies. This approach entails measuring the number of minutes that nurses, nurse aides, therapists, and therapist aides spend in caring for nursing home patients, and using these measures as the dependent variable in the RUG-III classification system. In these studies, field researchers measure staff time use over the course of a week for individuals in the study samples. A key advantage of this staff time measurement method is that it captures the resource consumption of nursing care at the individual person level. In analyses leading to the staff time weights of the existing RUG-III classification system, for example, the amount of explained variance in staff time was approximately 50% (Maxwell, et al. 2003). The principal limitations of such studies, however, are that they require considerable time and money to conduct, and, perforce, can only reasonably study a limited number of patients.

In contrast to the SNF PPS, Medicare's other PPSs establish payment weights using routinely collected administrative data, such as charges for individual patients' care. This traditional method has several advantages: most notably, the ability to use the entire Medicare patient population as a data source in developing both the payment weights and the classification system; the ability to capture all resource consumption, including prescription drug costs; and the ability to be easily recalibrated with more recent cost and charge data once the PPS is in place. Because the nursing component of resource use is only available at the facility level, this approach is most useful when nursing care is a relatively small proportion of total resource use. In the case of Medicare SNF services, however, nursing care is a large share of total resource use. Furthermore, measures of average nursing cost per patient obtained from administrative data are generally more indicative of non-Medicare rather than Medicare patients, since Medicare patients generally constitute only a small proportion of total patients in nursing homes.

In sum, the three primary areas of interest listed above were our primary focus for refinement research efforts. The growth in NTA costs highlight the importance of refining the SNF PPS' ability to capture the costs of patients with high NTA service needs, including drug and respiratory therapy. At the same time, the large proportion of patients in the rehabilitation therapy category raises questions on whether the SNF PPS encourages provision of too much therapy, particularly in light of the absence of clinical knowledge on the appropriate amount of therapy that should be provided to achieve optimal outcomes. The potential need for rebasing the SNF PPS also encourages a focus on making greater use of routinely collected information on Medicare patients—including SNF claims, patient assessments, and claims from SNF patients' prior hospital stays.

III. Data and Methods

This chapter briefly describes our data sources and methodological approaches. More detailed information can be found in the Chapter III Appendix. First, we describe the data sources and creation of the 2001 SNF analysis file. Second, we address the selection of units of observation, and describe our approach for dealing with timing inconsistencies between claims data on costs and MDS data on patient characteristics. Third, we briefly describe the dependent and explanatory variables examined. More details are included in the subsequent chapters that address the specific patient classification models. Fourth, we highlight analytical issues that have implications for all classification models. Finally, we highlight our criteria for statistical evaluation of the different models.

A. Data Sources

The principal data source for this study is the 2001 Data Analysis PRO (DataPRO) file. This file consists of Medicare SNF stays that are linked with minimum data set (MDS) assessments, as well as information merged from the qualifying (or prior) acute care hospital stays. DataPRO was initially created to facilitate a medical review process of Medicare SNF patients and contains nationally representative information on Medicare SNF stays. About two million such stays occur each year. The DataPRO file contains information on patients' primary and secondary diagnoses, cognitive status, functional status, nutritional status, periods of service use, procedures, incurred provider charges, and other information from the SNF and prior hospital stays. We enhanced the DataPRO stay records with additional information from

Medicare claims, including charges for specific types of services (e.g., respiratory therapy, prescription medicine) used during the SNF stay and the prior hospital stay. Medicare claims data from the qualifying hospital stay provided diagnosis codes, as well as indicators and charges for a number of services. To the extent that these services must be continued in the SNF and that some services, such as intensive care unit (ICU) stays, by definition identify sicker patients, information from the qualifying hospital is likely to be highly predictive of costs for the SNF stay.

We also used data from the cost reports that Medicare-participating SNFs submit annually to fiscal intermediaries. Among other things, these reports itemize Medicare-related costs and charges for routine and ancillary services, and costs of capital. We used SNF cost report data to derive routine costs and the ancillary service cost-to-charge ratios (CCRs) used to convert ancillary service charges from claims to estimated costs. Because charges for routine services (e.g., nursing, accommodations) are not generally differentiated on patients' claims in the same facility, we assigned the per diem routine costs of a patient based on each SNF's reported routine service costs in the cost reports.

The analysis file for this study is a combination of the DataPRO file, which we enhanced with additional variables, and information derived from Medicare cost reports. Because both the claims and cost report information are necessary to allow us to estimate ancillary costs from claims, the analysis file contains 2001 SNF stays that: (a) have cleanly matched MDS information on patient characteristics; (b) come from facilities with 2001 cost reports; (c) have consistent claims information on SNF and hospital services not captured by DataPRO; and

(d) have internally consistent information from the multiple data sources. Appendix III-1 provides further detail regarding the construction of the analysis file.

B. Samples and "Base Case" Facility and Stay Data Exclusions

After data exclusions, the original analysis file contains approximately 1.7 million Medicare SNF stays in 2001, with about 600 variables. Because of the enormous size of this file, we selected a 10% random sample of stays for the purpose of developing case-mix classification models. We refer to this as the "test sample."

Table III.1 Number of Cases Remaining after Specific Initial Edits

Reasons for Exclusion of Stays	Remaining Numbers
Total SNF stays in 2001 DataPRO	2,114,797
1. "Integrity Problems and Non-Medicare coverage"	1,900,036
2. Unavailability of MDS assessments	1,778,059
3. Mismatch of the DataPRO stays and supplemental claims data	1,768,761
4. Matching with cost report information and wage-index file	1,722,987
5. Other exclusions	1,709,736

A second sample was drawn for validation and facility-level analyses. A random 10% sample of facilities was identified and all stays in those facilities were included in the second sample. We refer to this sample as the "validation sample."

Application of cost to charge ratios

Because of our interest in examining different levels of aggregation for ancillary services, information was "built up" to the desired levels. For example, the sum of department specific costs for NTA services and the sum of department specific charges for NTA services were used to construct the cost to charge ratio for all NTA services (rather than using the "total" cost and charge fields on the cost reports).

Because a major focus of this study aims to develop patient classification via NTA costs, we also disaggregated NTAs into more detailed components: (1) drugs, (2) respiratory therapy, and (3) all other NTAs. In the analysis, we examine the CCRs for those three components of NTAs, as well as for: (4) total NTAs, (5) rehabilitation therapy, and (6) total ancillary services.

The number of stays in each sample before and after the base-case exclusions is given in Table III.2.

Table III.2: Number of SNF Stay Observations in the 10% Stay and 10% Facility/Stay Analysis Samples

	10% Stay File (N)	10% Facility/Stay File (N)
Before exclusions	170,774	177,960
After facility exclusions	167,113	174,263
After stay exclusions	163,738	170,783

C. Analysis

Because preliminary analyses indicated that a much higher proportion of the variance in costs per diem could be explained than of costs per stay, we focused our analysis on costs per diem. In several analyses, we also compared explanatory models in terms of their ability to explain charges per day relative to costs per day. We believed that this type of comparison would be particularly informative, given the wide range of CCRs calculated from the cost reports. In all our analyses, we adjusted the SNF labor share of the cost and charge variables by the area wage index CMS uses in its hospital and SNF payment systems.

D. Explanatory Variables

For the analysis of SNF costs and charges, we obtained explanatory variable data from four sources: the DataPRO stay files, the MDS, additional data from the SNF claims, and prior or qualifying hospital claims. The potential explanatory variables included demographics, primary diagnoses, service indicators, comorbid conditions (i.e., secondary diagnoses), functional status indicators, and facility characteristics. Some variables, such as clinical diagnoses, were obtained from multiple sources.

Table III.3 illustrates the types of variables that we examined. The average age of SNF patients was eighty-years-old and two-thirds of the patients were female. Congestive heart failure, diabetes, and chronic obstructive pulmonary disease were common diagnoses among SNF patients, with each diagnosis recorded in approximately one quarter of the cases. The patients were also medically complex; nearly 60% of the sample possessed five or more comorbidities; more than one-third (35.7%) of the patients were cognitively impaired; and about 50% of the sample had one or more dependencies in Activities of Daily Living (ADLs). In 2001, the average SNF stay was 24.3 days, and the average duration of the prior hospital stay was 9.2 days. More than one-fifth (21.1%) of the SNF patients had been in intensive care units (ICU's) in their prior hospital stay. Most (78.7%) of the SNF stays were in facilities located in metropolitan statistical areas (MSAs) and 23.4% were in hospital-based facilities.

Table III.3: Average Characteristics of SNF Patients, 2001

Age	80.0 years
Female	65.9%
Congestive Heart Failure	26.6%
Diabetes	23.7%
Chronic Obstructive Pulmonary Disease (COPD)	23.6%
5 + comorbidities	58.3%
Cognitive Impairment	35.7%
Totally Dependent in Function	
1 + ADL	49.3%
3 + ADL	28.7%
Duration of Stay	24.3 days
Duration of Prior Hospital Stay	9.2 days
ICU use in Prior Hospital Stay	21.1%
In MSA SNF	78.7%
In Hospital-Based SNF	23.4%

E. Analytical Issues

A number of important analytical issues surfaced over the course of our research that have implications for all the specific approaches we explored:

Coding of Diagnoses

We discovered in prior research (Urban Institute 2004) that it is advantageous to analyze the relationship between patient diagnoses from claims data and SNF costs using projections from the prior hospital stay rather than the SNF stay. There are two important reasons for this conclusion. First, freestanding SNFs often fail to code secondary diagnoses on the claims, since secondary diagnoses do not influence Medicare SNF payment. On the contrary, hospital-based SNFs seem to follow the inpatient pattern of coding such diagnoses more frequently, even though these codings do not affect their SNF payments either. Because of this difference, measuring comorbidities with SNF diagnoses would "short change" freestanding SNFs. Using diagnosis information from the prior hospital stays makes freestanding and hospital-based SNFs comparable in coding practices regarding secondary diagnoses. Second, hospital-based SNFs often (one-quarter of cases) record the code "rehabilitation" as the primary diagnosis for the SNF stay, whereas freestanding SNFs rarely use this code. In analyses of costs, therefore, the prevalent code of "rehabilitation" would function as a proxy for hospital-based SNF status, thereby introducing a distorting "facility characteristic" determinant of patient costs.

Ambiguity in the Meaning of MDS Service Use

Some of the MDS questions are posed in terms of "the past 14 days." That is to say, for certain assessments (e.g., 5 day), service could have actually been delivered during the prior hospital stay rather than the SNF stay. For example, we found that about half the SNF patients had received intravenous (IV) medications in the past fourteen days. Such a high proportion is simply implausible for patients during an SNF stay—a hypothesis that was confirmed when our comparison of the MDS responses on IV medications for Medicare SNF and prior hospital stay claims found a very large discrepancy between the two sources. Because we cannot necessarily attribute use of particular services recorded from the MDS to actual use in the SNF, we define IV medication use as occurring in the SNF only if both MDS *and* SNF claims indicated some IV-related activity.

The Small Size of the Medicare Share in Most SNFs

Because Medicare revenues are only a small share of total revenues for freestanding SNFs, the overall costs and practice patterns of those SNFs tend to reflect those of non-Medicare patients. To the extent that care provision between Medicare and non-Medicare differ, the amount or type of services provided to Medicare patients is likely to be constrained by facility resources more readily available to non-Medicare patients. Thus, for a given set of patient characteristics, associated costs may have a considerable range more because of facility characteristics than because of patient need.

F. Statistical Evaluation of Models

This section describes the general features of the models used in each patient classification approach and the minimum common set of statistics that are reported and used to evaluate each approach. Further discussion and technical detail is provided in Appendix III-2.

Patient Stay-Level Models and Statistics

Classification systems used in a PPS should be able to account for a reasonably high proportion of the predictable variation in a provider's patient care costs due to clinically meaningful differences in patient characteristics. To the extent that a classification system does not sufficiently meet this goal, provider incentives to select patients according to risk may increase. Predicting variation in costs that is due to clinically inappropriate variation (e.g., provider inefficiencies or regional practice patterns not related to best practices) is not desirable. In other words, higher predictability of cost variation is generally considered better, as long as it is the "right kind" of variation. The unit of payment currently used for SNFs under Medicare is based on per-diem patent care cost. Thus, a classification system should explain sufficient appropriate variation in costs per day at the person level.

The general form of the stay-level model is:

$f(\textit{wage-adjusted cost per day}_i) = g(\textit{patient classification system variables}_i)$,

where f and g are functions and i denotes a patient stay. We leave the functional form arbitrary because models may be estimated in a variety of ways, including linear regression on costs, linear regression on the log of costs, two-step regression, and classification and regression tree analysis (CART).

CART (Classification and Regression Tree Analysis)

Traditional regression models have disadvantages. They require assumptions about the parametric functional relationship between resource use and explanatory variables, such as patient age. For instance, resource use may not change much for stays by patients ages 65-80 years, but may increase drastically for stays by patients over 80-years-old, a non-parametric relationship not captured in the regression model. They also do not identify interactions between explanatory variables, unless specified *a priori*. For example, a person with anemia may only require expensive drugs in the presence of renal failure, an interaction. In estimating resource use for SNF stays, both issues may be of concern.

A regression tree approach allows these types of relationships to be more effectively modeled by searching for partitions or split points in the variables. In the age example just mentioned, the model would look for the best partition (in terms of minimizing unexplained variation) of age, in this case, 80-years-old. A CART model searches among all variables and split points recursively, creating ever more complex interactions. It first selects a break in the values of a variable that results in the highest amount of variance explained. It then progresses in this manner to continue making breaks in the variables until no further gain is obtained in variance explanation. The output of a CART analysis is a "tree" with multiple branches, and "endpoints" or buckets representing the interaction of the variables that led to those endpoints.

The drawback of regression trees is the tendency to 'overfit' the data. While the splits and subgroups may be informative regarding interactions between explanatory variables, the high order interactions produced through recursive splitting may not fit the population of interest, but instead be an artifact of the sample.

R-squared

The R-squared statistic reported for each model tells us what share of the variation in the dependent variable (e.g., costs per day) can be accounted for by the patient classification system. In linear regression models, the usual R-squared is reported (explained sum of squares / total sum of squares). For two-step or uncentered models (where the average predicted cost is not equal to the average actual cost) and for out-of-sample predictions, we report the R-squared as computed by a regression of the actual cost on the predicted cost. In comparing models with the same dependent variable but different sets of explanatory variables, so long as the cost variation being explained is actually attributable to clinically appropriate patient characteristics, the classification system with the highest R-squared is preferred, all else equal.

R-squared is an overall, or summary, measure of predictive ability. Two models with similar R-squared could differ in how well they predict cost for high cost patients versus low cost patients. We can obtain more detail from measuring sensitivity.

Sensitivity

Sensitivity measures how well the classification system can correctly predict high cost cases to be high cost. We define high cost cases for this purpose as cases in the highest 10% of costs. Sensitivity is calculated as the percent of cases in the highest 10% of actual costs that are also in the highest 10% of predicted costs. As such, the sensitivity measures tell us the probability that the most expensive stays will be paid at the highest payment rates under a given classification system. All else equal, a more sensitive classification system will be less likely to create incentives for providers to select against the most costly patients.

Standard Deviation of the Relative Weights

A patient classification system yields a set of predicted costs that depend on patient characteristics. Relative weights (also called relative values) measure the costliness of a patient type relative to the average patient. Relative weights are constructed by dividing the predicted cost for a particular type of patient by the average actual cost (or average predicted cost).[5] The relative weights have a mean of 1.0 in the sample from which they are derived. They normally have an average close to, but not exactly, 1.0 when applied to broad groups out-of-sample. For each of the base case models, we report the standard deviation of the relative weights, which provides an indication of the extent of variability of the payment rates that will result from a particular payment system.

Validation Out-of-Sample

The base case models are estimated using the test sample. For out-of-sample validation, the results of these models are then applied to the validation sample. The evaluative statistics are reported for both the test and validation samples for each model. Out-of-sample validation

[5] The average actual cost and average predicted cost are the same in one-stage regression models that contain a constant term.

mimics the way classification systems are used in practice, whereby payment rates developed from older data are used for determining current payments. Large differences between test sample and validation sample results may indicate substantial over-fitting of models to the data in the test sample.[6]

Facility-Level Models

Effective classification systems should also be able to account for variation in costs across facilities that are attributable to variation in average patient case-mix characteristics. Facilities may have high average costs due to a more severe or complex patient mix, perhaps because they are well equipped to treat such patients. Facilities with a more severe case-mix would be penalized if the patient classification system does not reflect these differences such that financial losses on high-cost patients sufficiently offset gains from low-cost patients (MedPAC 1999). Differences in facility payment levels due to case-mix differences should both explain and be proportional to facility differences in average costs per patient day in SNFs.

We estimate facility-level cost models to address these issues, examining how facility costs vary with case-mix-adjusted payments and other facility characteristics. We use the validation sample for this purpose. Stay-level costs are averaged, weighting by Medicare covered patient days, to construct the average cost per patient day in each facility.

[6] Because the statistics presented for the validation sample provide a more realistic appraisal of the statistical properties of the classification system, we rely on the validation sample results in comparing the different approaches.

Facility Case-Mix Indices

Facility case-mix indices (CMIs) summarize the costliness of a facility's patient case-mix as captured by a particular classification system. CMIs are constructed by averaging the (out-of-sample predicted) relative weights of the stays within each facility in the validation sample. The average CMIs are weighted by the number of covered days for each stay. At the facility level, CMIs are recentered so that the means of the CMIs across all facilities are equal to 1.0. We estimate facility-level models by regressing the log of facility average costs per day (unadjusted for area wages) on a constant, the log of the CMI, and, possibly, on a set of control variables, which varies according to the model specification. We estimate and report three specifications using the base case from each classification approach:

CMI-only model:

$$\log(\textit{Facility cost per day}) = \beta_0 + \beta_1 * \log(\textit{CMI}) + \varepsilon_1$$

Payment model:

$$\log(\textit{Facility cost per day}) = \gamma_0 + \gamma_1 * \log(\textit{CMI}) + \gamma_2 * \log(\textit{WIF}) + \gamma_3 * \textit{Urban} + \varepsilon_2$$

Fully-specified model:

$$\log(\textit{Facility cost per day}) = \delta_0 + \delta_1 * \log(\textit{CMI}) + \delta_2 * \log(\textit{WIF}) + \delta_3 * \textit{Urban} + X\delta_4 + \varepsilon_3.$$

The CMI-only model contains no control variables except for a constant term. The payment model adds the log of the area wage index factor (WIF) and an urban facility indicator as control variables, because these two variables are currently accepted as valid facility characteristics for which SNF payments should be adjusted. The "fully-specified" model adds control variables, denoted by *X*, for number of beds (0 to 9 beds, 10 to 19 beds, 20 to 49 beds, 50 to 99 beds, with 100+ beds as the excluded category); being a hospital-based facility; ownership type (voluntary/nonprofit, for profit-other, and government, with for profit-chain as the excluded category); and percent of SNF patient days paid by Medicare. These additional variables are expected to correlate with facility costs and may be correlated with the CMIs and unobserved aspects of patient case-mix; however, they are characteristics that are not considered appropriate to use in adjusting SNF costs *per se*.

Facility-Level Model R-squared

The R-squared of the CMI-only model indicates the maximum share of variation in (log) average SNF costs per day that can be attributed to facility case-mix as measured by the particular classification system. This measure can be compared across different classification systems for the same cost component. It is also informative to examine how the R-squared value changes as additional explanatory variables are added. The change in R-squared from the CMI-only model to the payment model summarizes the potential contributions of further payment adjustments for area wage levels and urban status. The change in R-squared from the payment model to the fully-specified model summarizes the unique contribution of the other facility characteristics that are considered inappropriate for payment adjustment in explaining SNF costs.

The CMI Coefficient and CMI Compression

The CMI coefficient (i.e., the regression coefficient on the log (CMI) variable in a facility-level model) measures the percent change in expected facility costs associated with a percent change in the CMI (and, thereby, its average relative payment rate per patient day due to case-mix). The ideal situation is for the CMI coefficient to be 1.0, indicating, for example, that a 10% higher CMI is associated with 10% higher costs. When the CMI is proportional to expected costs, facilities are paid in proportion to the higher cost associated with a higher CMI.

The situation that arises with a CMI coefficient that is greater than 1.0 is called "CMI compression." If, for example, the estimated CMI coefficient is 1.25, a facility whose CMI and average payments are 10% higher than average would tend to have costs that are 12.5% higher than average. Costs tend to rise (and fall) more than payments, so that high case-mix facilities tend to be underpaid and low case-mix facilities tend to be overpaid. The term compression is used because the distribution of case-mix-adjusted payments is more narrow, or compressed, than the distribution of expected costs associated with facility case-mix. A classification system that exhibits CMI compression would tend to penalize facilities that have a more severe case-mix and creates incentives to avoid high cost patients (Cotterill 1986). Alternatively, a CMI coefficient that is less than 1.0 would indicate what is called "CMI decompression."[7] Although decompression is a less common finding in the literature, it arises in some of the results we present below.

We report the CMI coefficient for each of the three specifications. In addition to the CMI, we report the standard error of the CMI coefficient, so that confidence intervals can be computed, and the p-value for the t-test of the null hypothesis that the CMI coefficient is equal to 1.0.[8]

---

[7] The term "decompression" could also be used to describe a procedure to adjust for compressed weights, but here we mean the opposite of compression.

[8] The CMI coefficient results for the CMI-only specification are the least relevant when we consider that payment adjustments are made for area wages and urban status. They are reported primarily for completeness, although they are also useful to illustrate whether wage and urban adjustments alter the relationship between the CMIs and costs.

Interpreting Compression Results for the Payment and Fully-Specified Models

Findings of CMI compression or decompression have somewhat different meanings and implications in the payment and fully-specified models. The CMI coefficients in the payment models quantify the relationship between facility case-mix and costs across facilities, holding area wages and urban status fixed (or as if an average cost adjustment has already been made for those two factors). The CMI coefficient tells us whether case-mix adjusted payments will flow to facilities proportionately to their expected costs.

The CMI coefficients in the fully-specified models measure the extent to which facilities of a given type would face incentives to risk-select patients on the basis of the characteristics included in the classification system. Facilities may be making or losing money overall. However, if costs are related proportionately to the CMI in the fully-specified models, facilities of a given type will not be making or losing more or less money depending on the case-mix characteristics of their patients. We provide more detail on the interpretation of the CMI coefficient results and potential sources of CMI compression in Appendix III-2.

If the fully-specified model were in fact a reasonably accurate overall model of the drivers of SNF costs, compression in the fully-specified models would cast serious doubt on the effectiveness of a classification system that is not otherwise modified with payment cushions, such as outlier payments. Decompression would also suggest serious problems that an outlier policy would not address.

IV. RUG-Based Approaches to Classifying NTA Costs

A. Aims and Background

Our research to develop a RUG-based approach to better account for variation in non-therapy ancillary (NTA) costs reexamines and extends prior research that focused on this approach. We reexamine prior efforts by Abt Associates (2000) and Fries (2003) using data from 2001 analysis files (Urban Institute 2001), and extend the research by exploring consolidation of RUG-III groups and by testing additional variables, particularly those that define services received by SNF patients when they are indicated on *both* MDS and claims data.

Research by Abt Associates (2000) identified a number of different strategies to enhance the ability of the RUG-III system to explain NTA costs. First, Abt Associates discovered that individuals who qualified for both the "rehabilitation therapy" and "extensive services" sets of categories in the RUG-III had NTA costs that were much higher than average. This finding led to the recommendation by Abt Associates to add fourteen categories to the RUG-III to represent the intersection of these two RUG-III categories. This new classification scheme is referred to as RUG-58.

The research by Abt Associates, as well as Fries (2003), led to the creation of indices based on specific MDS items that could be appended to the RUG-III system. Abt Associates explored the variables in the MDS for patient characteristics associated with variations in NTA costs. After identifying variables that were statistically associated with NTA costs, Abt Associates convened clinical panels to consider which of the variables were least likely to be "gamed" or lead to quality of care problems if they were used to determine payment levels. In the process, they explored changes in definitions that could significantly reduce such problems. For example, their definition of oxygen administration was restricted to having received oxygen due to particular diagnoses or symptoms. This process resulted in the identification of twelve variables from the MDS, including both patient conditions and service use. Two alternative indices were created using these variables. The first index was based on a regression analysis in which NTA costs were regressed on the twelve conditions and services. An individual's score would essentially be the sum of the conditions or services he or she used, "weighted" by the coefficients in the regression model. This index was called the weighted index model (WIM). The second index was simply a count of the number of the twelve conditions or services an individual patient possessed, and was labeled the un-weighted index model (UWIM). In effect, the more conditions or services reported for an individual, the higher would be his or her NTA costs.

In exploratory analyses with 1999 SNF cost and MDS data, we found that the RUG-58 explains about 6% of the variation in log NTA costs, while the WIM-type indices increase the amount of explained variance to 10%. These percentages are slightly lower than the variation explanation levels of NTA costs found by Abt Associates (2000) in their original research. A full comparison of our findings with those of prior research on the RUG-58 and the WIM-type indices are presented in materials we prepared for a Technical Advisory Panel (TAP) meeting in May, 2003 (Fries 2003).

This chapter presents our recent findings on RUG-based research approaches using the 2001 analysis file. We re-examine the RUG-58 and WIM-related indices, as constructed by Abt

Associates, and explore variants based on the use of additional variables reflecting patient conditions and services used. We also examine the effects of consolidating RUG-58 groups, for example, by collapsing the twenty-eight rehabilitation therapy groups into two categories according to whether the case also qualifies for the extensive services category. Based on much testing of multiple models, we identify two "preferred models" for addressing NTA services in the context of RUG-based approaches: RUG-58 and the SIM model.

B. Data and Methods

Exploratory Findings

The primary data sources for our analyses were the test and validation files (both described in Chapter III). We also used a 1% SNF stay file to systematically explore variables that were correlated with total NTA costs.[9] We selected particular variables for exploration when the value of their t-statistic was notably higher than those of other variables. For groups of variables (e.g., diagnosis) and continuous measures we examined the R-squared statistics from a regression including only those variables.

We explored two types of variables from the SNF stay files. The first type consisted of variables derived directly from the data sources—MDS assessments, DataPRO SNF stays, and additional revenue center variables from claims that corresponded to the DataPRO SNF stays. The second type consisted of the intersection of variables from MDS assessments with revenue

---

[9] In each file, in addition to dropping the cases described in the previous chapter, we eliminated a small number of cases (e.g., 0.3% of the cases in the validation file) for which no RUG category was assigned.

center indicators from the claims data: Presence of a service was activated only when indicated both by questions on the MDS assessments and by SNF claims. The intersection variables were initially constructed to address the problem that some service use recorded on the MDS reflects services received during the prior hospital stay (as discussed in Chapter III).

The most important finding of our exploratory research showed that measuring IV medication or respiratory care based on both MDS assessments and claims dramatically improve our ability to distinguish the high cost cases, compared to measuring service use with either source independently. Table IV.1 shows the average wage-adjusted NTA cost per diem for groups of stays with and without IV medication (panel 1) and with and without respiratory/pulmonary care (panel 2).[10] Within each panel, the stays are separated according to whether the information regarding these treatments were obtained from the MDS, the claims, or both. For both IV medication and respiratory care, stays with service use indicated by both MDS assessments and the claims have much higher NTA costs than stays with service use indicated by either source alone.

Table IV.1 also shows that higher NTA costs are found in stays with both claims and MDS indications of IV therapy, regardless of which MDS assessment is examined. For example, at the thirty-day assessment, considerably higher per diem NTA costs are observed for stays with MDS assessments showing IV therapy and claims showing IV solution or therapy ($137) than for those with either the MDS indicator alone ($80) or claims alone ($62). This is similar in size

[10] In the test sample, of all stays with an indication of IV therapy use, 17.1% had IV indicated by both the MDS and claims, 3.7% by claims only, and 79.2% by the MDS alone. For respiratory care, 19.5% had an indication on the MDS and claims, 28% on claims only, and 52.6% on the MDS alone

to the differential observed for the 5- and 14-day assessments. A parallel pattern can be seen in the interaction of MDS indicators of respiratory need and the claims indicating such use.

Table IV.1: Average Wage-Adjusted NTA Stay Cost Per Diem by Presence of MDS and Claim Report, by Assessment Number

Assessment	Neither MDS or claim	MDS only	Claim only	MDS and claim	N
IV Medication*					
5 day	45.2	64.9	93.2	158	170,117
14 day	35.0	66.2	71.9	141	101,754
30 day	31.6	79.7	61.6	137	48,521
60 day	31.2	71.3	58.8	127	14,554
90 day	31.3	74.3	59.5	134	5,256
Respiratory Problem**					
5 day	50.9	72.1	124	163	131,256
14 day	38.7	51.7	98.8	136	101,754
30 day	34.0	46.8	75.1	120	48,521
60 day	31.0	44.5	60.2	125	14,554
90 day	30.4	48.2	55.3	139	5,256

*Indicates IV Medication reported on MDS for the given assessment and/or a claim for IV therapy or solution for the stay.
**Indicates tracheotomy, vent or oxygen (with conditions) reported on MDS for the given assessment and a claim for pulmonary or respiratory for the stay.
Source: Validation File extracted from 2001 DataPRO.

These findings suggest that the cases indicated as needing IV medication or respiratory care based on both the MDS assessments and the claims data are different in some respects from those indicated by either source alone.   For the assessments that occurred in the first fourteen days of the stay, use of the interaction excludes cases that only received services prior to the SNF stay. Recall that the MDS defines service use as services used in the past fourteen days, which may also include the hospital stay. The continuation of the pattern to later assessments suggests that the intersection reflects an intensity of service use, possibly leading to recording by both sources. An additional explanation of this difference is that error in measurement is heightened if each source is used independently— including some cases without the condition and excluding some with the condition. Requiring use of both indicators to define the condition may limit the extent of such errors.

It is important to note that a respiratory service indicator on claims suggests care provision by respiratory therapists. It is appropriate, however, for nurses who are billed under routine costs also to provide respiratory therapy in SNFs. To operationalize this variable in an eventual payment system, we need to consider approaches that do not discourage provision of respiratory services by nurses, within the scope of their expertise, or provision of more specialized procedures by respiratory therapists when needed.

Models of NTA Costs and Charges

We evaluated three sets of models of NTA costs: RUG-58, WIM, and the Service Index Model (SIM).

Construction of the RUG-58 and WIM variables followed the procedures developed by Abt Associates (2000). In addition, we created a new classification, RUG-58G, in which the fifty-eight RUG categories are grouped into eight broader categories. (The G refers to “grouped”.) These categories include: (1) rehabilitation with extensive services; (2) rehabilitation without extensive services; (3) extensive services without rehabilitation; (4) special care; (5) clinically complex; (6) impaired cognition; (7) behavior problems only; and (8) reduced physical function.

A new index that we developed, which we refer to as the *Service Index Model* (SIM), is a modification of the WIM index. The SIM model mimics the basic approach of the WIM model with several important distinctions. First, we took advantage of the sizable increase in predictive power obtained by defining IV medication use and respiratory care using the intersections of MDS measures with indicators from claims. Second, we used several other measures of the patient condition and history found in our exploratory analyses that add predictive value to the model and that we believe are difficult to game. These variables are: (a) age, (b) absence of infection, and (c) major diagnostic category (MDC) for respiratory conditions.

We estimated the index models using regression analysis on the stay-level data. Per diem NTA costs were regressed on patient characteristics. The resulting coefficients indicate the effect of a given characteristic on NTA costs while controlling for the other characteristics in the model. Using this basic approach, we examined NTA costs and charges in levels (rather than logs). About 2% of stays have zero NTA costs and charges; these were included in the model without adjustment.

C. Findings

We present our findings in five sections. First, we highlight the findings from our exploratory analysis of patient characteristics and services used by focusing on variables used in RUG-III and the WIM and SIM indices. Second, we compare the extent to which different RUG-related models explain variation in NTA costs. Third, we illustrate how payment cells for NTA can be created from an index model. Fourth, we discuss the aggregation of person-level data to create facility-level case-mix indices (CMIs). Finally, we highlight person-level and facility-level statistics on two 'preferred models' to address NTA costs under the aims of RUG-based approaches.

Bivariate Relationships Between Model Variables and NTA Costs

We first examined RUG-related categories (RUG-III and RUG-58) and their relationship to NTA costs. Our general finding was that only two of the individual categories in either RUG-III or RUG-58 independently explain over 1% of variance in NTA costs (Extensive Services, levels 2 and 3).[11] This finding is consistent with prior research indicating that RUG-III, as a whole, explains only about 5% of the variation in NTA costs.

Next, we examined the variables from the WIM and SIM models. Table IV.2 presents the coefficients and R-squared of bivariate regressions of per diem NTA costs on the individual variables used in the WIM; Table IV.3 presents the same for the variables used in the SIM.

Some of the individual WIM variables in Table IV.2 predict costs well, with IV medication use explaining around 7% and respiratory service needs explaining around 2.6% of the variance in NTA costs. As seen in Table IV.3, the effectiveness of those variables, when restricted to instances with claims indication as in the SIM model, is considerably greater. In the SIM model, IV medication use explains 13%, while respiratory-related needs explain 6% of the variation in NTA costs.

Although the variables used in the SIM do a substantially better job of predicting NTA costs, it might be difficult in practice to use them for payment. The claims and the MDS data do

[11] In addition, a regression of cost on the RUG-58 rehabilitation therapy category, RVB, has an R-squared of exactly 1%.

not come in at the same time, and there is currently no mechanism for linking them on a claim-by-claim basis. However, it is clear that if the measure of service use in the MDS can be redefined to capture the characteristics of stays with both MDS and claims reporting service use, the new measure could dramatically improve the ability of a classification system to predict NTA costs.

Table IV.2: Coefficients from Bivariate Regressions of NTA Cost Per Diem on WIM Variables, Estimated Using 2001 DataPRO Test File

Variable	Mean	Regression of NTA per diem on Variable	
		Coefficient	R-squared
IV medication (MDS)	0.402	53.8	0.073
Oxygen and either pneumonia or respiratory infection (RI) with fever, or pneumonia or RI, COPD, CHF, CAD with SOB	0.131	43.6	0.026
Parenteral/IV with >76% total calories	0.011	158	0.032
Suctioning	0.025	77.6	0.018
Tracheotomy care	0.010	91.4	0.011
Pneumonia	0.126	33.7	0.015
Tube feeding with >76% total calories	0.069	34.1	0.009
RI	0.037	27.0	0.003
Application of dressing with/without topical medication, and presence of ulcers or other skin lesions/wounds	0.077	16.3	0.002
Skin wound/ulcer care (MDS)	0.387	16.0	0.007
Stage 4 pressure ulcer (MDS)	0.045	26.0	0.003

N=163,386

Table IV.3: Coefficients from Bivariate Regressions of NTA Cost Per Diem on SIM Variables, Estimated Using 2001 DataPRO Test File

Variable	Mean	Regression of NTA per diem on Variable	
		Coefficient	R-squared
IV medication (MDS) *and* claim for IV therapy or solution	0.081	121	0.132
Oxygen (linked to conditions) or tracheotomy care or ventilator *and* claim for respiratory or pulmonary	0.041	107	0.055
Age	80.0	-1.38	0.023
Respiratory disease in SNF stay	0.294	29.4	0.023
No infection (MDS)	0.414	-17.4	0.009

N=163,386

Multivariate Models

In Table IV.4, we present findings from a multivariate regression model of per diem NTA costs on the SIM variables. We find that the effect of IV medications is quite large, with a $107 difference associated with the interaction of IV medication and claims. This effect is roughly 2.5 times that found without including claims information in the definition. The effect of respiratory therapy—oxygen/tracheotomy/ventilation—intersected with claims indicators of these services is $75, which is larger than the distinct effects of either oxygen use or having had a tracheotomy. The share of variance in NTA costs explained by the SIM model is 19%. Though not shown, this result is considerably larger than the 13% R-squared associated with using the WIM variables alone. The model correctly predicts 45% of high cost cases. A model of charges shows similar patterns, with a more dramatic increase in R-squared.

Table IV.4: SIM Model for Predicting Adjusted NTA Costs Per Diem – Estimated Using 2001 DataPRO Test File

Variable	Coefficient	Test Statistics
IV medication (MDS) *and* claim for IV therapy or solution	107	53.1
Oxygen (linked to conditions) or tracheotomy care or ventilator *and* claim for respiratory or pulmonary	75.0	35.7
Age	-1.08	-37.6
Respiratory disease in SNF	15.7	27.7
No infection (MDS)	-7.50	-12.5
Constant	136	55.7
R-squared (all variables included)		0.189
% Predicted >=90th percentile given actual >= 90th percentile		0.453
Standard deviation of relative weight		0.300

N=163,386, test statistics are adjusted for clustering by facility.

Highlighting Variance Explanation of Different Models

The three sets of models tested – RUG-58, WIM, and SIM – vary greatly in their ability to explain NTA costs.[12] RUG-III explains 4 – 6% of the variation in NTA costs, while RUG-58 explains 9.5% of the variance.[13] The WIM variables alone explain 12.5% of the variation in NTA costs and the SIM variables alone explain 18.9% of the variance. Together with the RUG-58 variables, WIM explains 13.9% of the variance and SIM explains 21.9%. These results are presented in Table IV.5 and discussed in detail below.

---

[12] Ability to explain variation in resource use is fundamental for a case-mix classification system to appropriately reimburse providers.

[13] These models are based on levels of costs. Estimating the models with either charges or log costs as the dependent variables has little effect on the relative R-squared of the various model specifications.

Table IV.5: Comparison of Predictive Models Incorporating RUG-III, RUG-58, WIM and SIM, Estimated Using 2001 DataPRO Test File

Explanatory Variables	R-squared	% Predicted >=90th percentile given actual >=90th percentile	Standard Deviation of Relative Weights
RUG-III nursing payments, 2001	0.035	0.258	0.303
RUG-III (new weights)	0.064	0.314	0.354
RUG-58	0.095	0.376	0.429
RUG-58G	0.077	0.255	0.388
RUG-58 defined with claims interaction	0.128	0.368	0.499
WIM variables	0.125	0.325	0.492
WIM variables and RUG-58	0.139	0.362	0.521
SIM variables	0.189	0.453	0.606
SIM variables and RUG-III	0.214	0.457	0.644
SIM variables and RUG-58	0.219	0.456	0.652
SIM variables and RUG-58G	0.212	0.454	0.641
6-category CART based on SIM and RUG-58G	0.184	0.442	0.599
SIM and WIM variables and RUG-58	0.233	0.465	0.673

N=163,386

As a point of reference, the payment weights for NTA (i.e., the 2001 nursing weights), explain only 3.5% of the variance of costs. When we recalculate the weights using a regression of NTA costs on indicators of the RUG-III categories, variance explanation improves to 6.4%. Interacting the rehabilitation therapy and extensive services categories to create the RUG-58 increases the variance explained to 9.5%. However, replacing the RUG-58 with RUG-58G (in which the RUG-58 groups were collapsed to eight categories) reduces the R-squared to 7.7%. Generally, these models show only a moderate ability to predict the high cost cases: at best, 37.6% of cases with costs in the top 10% are predicted as having high costs. Adding the WIM to the RUG-58 increases the variance explained to 13.9% but has little effect on the share of high cost cases predicted. These results essentially represent a re-examination of prior research to refine the RUG classification system for NTA costs.

Table IV.5 also presents findings from our regression models using the SIM variables. These models considerably increase the share of NTA variance explained relative to the WIM models. For example, the SIM variables alone explain 18.9% of the variation in NTA costs; in combination with RUG-58, they explain 22% of the variance. Interestingly, substitution of the eight-category RUG-58G into the model produces nearly the same variance explanation as when SIM is included with the full set of RUG-58 groups.[14] Reduction of SIM and RUG-58G to only six mutually exclusive groups (described in the next subsection), explains 18.4% of the variation. In addition to the increase in R-squared, we find that the SIM models predict 45 - 46% of high cost cases to be high cost, as compared with 36% using the WIM model.

Finally, we experimented with redefining RUG-58 to incorporate the argument behind the SIM model. That is, we counted cases as having used IV medication and respiratory services only if a corresponding SNF claim was observed. We then reassigned cases to RUG-58 categories. This reassignment leads to an R-squared of 13%, which is an improvement on the original definition of RUG-58. However, the model shows little advantage relative to SIM itself.

In sum, the results in Table IV.5 indicate general bands of variance explanation of the models we tested. The basic RUG-58 model explains around 9% of the variance in NTA costs, a large improvement over the RUG-III currently in effect for the SNF PPS. The WIM index raises the R-squared to the 13%-14% level. The SIM index raises the R-squared to about 20%.

---

[14] Although models using the full and 1-digit versions of the RUG-58 model yield similar R-squared statistics, formal tests reject the premise that the RUG-58 coefficients are identical within each of the 1-digit categories

Interestingly, more parsimonious models within each band obtain nearly the same share of NTA cost variance explained as the more complicated models. For instance, the model combining the SIM indicators and RUG-58G results in an amount of NTA cost variance explained (21%) that is only modestly lower than more complex models that include the full RUG-58.

Application of SIM Models to Create Payment Cells

In this section, we use the SIM model to illustrate how such index models can be used to create a small number of payment cells for NTA services. We entered the SIM and RUG-58G variables into a classification and regression tree (CART) analysis. (See Chapter III's Data and Methodology section for a description and discussion of CART.) Results of our CART analysis are presented in Table IV.6. This model has six buckets. These were defined by CART and used as predictor variables in regression models. The R-squared (0.184) is fairly close to the R-squared from the linear regression model including both the collapsed RUG-58 categories and the SIM variables (0.212).

The six buckets can function as the payment categories for NTA services. For example, the lowest payment category is $39.50; the individuals for whom this amount would be paid are characterized by: (a) being in the rehabilitation category (but not also in extensive services), (b) not receiving respiratory-related services, and (c) not receiving IV medications. About half (51%) of the SNF stays belong to this NTA payment category. In contrast, the highest paying category is $247; the individuals for whom this amount would be paid are characterized by: (a) receiving respiratory-related services and (b) receiving IV medications. This category accounts for only 1% of SNF stays.

Table IV.6: Groupings and Payments from CART Model Estimated Using 2001 DataPRO Test File

Group	IV Medication	Respiratory Services	Rehab therapy but not also in Extensive Services category for RUG-58	Payment	Proportion of Stays
1	No	No	Yes	$39.5	0.511
2	No	Yes	Yes	112.0	0.011
3	No	No	No	66.5	0.377
4	No	Yes	No	152.0	0.019
5	Yes	No	All Cases*	164.0	0.071
6	Yes	Yes	All Cases*	247.0	0.011

N=163,386
*Includes all cases with indicated IV medication or respiratory services, whether or not in rehab therapy

D. Implications at the Facility Level

To evaluate the ability of each classification scheme to reimburse costs across a range of patient difficulty, we employ the CMI methodology described in Chapter III to estimate facility-level regressions of the log of per diem NTA costs on the log of the CMI. (See Chapter III for details regarding the construction and interpretation of the CMI.)

We evaluate the performance of RUG-III and two options among the alternatives, RUG-58 and SIM combined with RUG58G.[15] We evaluate the various classification schemes first using models based on costs and then charges to define the weights and CMI. Prior to controlling for differences across facilities, the cost-based RUG-III and RUG-58 models have coefficients on the CMI of close to one. However, after controlling for differences across facilities, only the cost-based SIM+RUG58G weights leads to a coefficient on the CMI of

[15]The RUG-58 is attractive because it is simple to implement and administer and its R^2 of 10% is a substantial improvement on both the current RUG-III nursing weights (4%) and the recalibrated weights for RUG-III categories (6%). The SIM+RUG58G is attractive because it builds on the RUG-58 model to explain about 20% of the variation in NTA costs.

approximately one. The cost-based weights from RUG-III and RUG58 and the charge-based weights from all three classification systems lead to coefficients on CMI of less than 1.0, implying that facilities with a low CMI have an incentive to attract more difficult cases to increase profits.

Models with Cost-Based Weights and Case-Mix Index

Table IV.7 presents comparative statistics on the RUG-58 and SIM+RUG58G, using NTA costs to calculate the payment weights and the resulting CMI. Comparable statistics for RUG-III using weights calibrated to NTA costs are included for reference. We focus discussion on RUG-58 and SIM+RUG58G models, grouping the statistics by person- and facility-level models.[16] This is followed by a corresponding discussion for models in which the payment weights and CMI are based on charges rather than costs.

[16] The person-level models are regressions of NTA costs on the classification measures. The facility-level models are regressions of log of facility-average per-diem NTA costs on the log of the CMI. The CMI is based on a cost regression.

Table IV.7: Comparison of Predictive Models of NTA Costs: RUG-III, RUG-58, and SIM+RUG58G[a]

Explanatory Variables	RUG-III (New Weights)	RUG-58	SIM+RUG58G[a]
Person-level models of per diem cost			
a. Test Sample			
R-squared	0.064	0.095	0.212
SD relative weights	0.354	0.429	0.641
Sensitivity[b]	0.314	0.376	0.454
b. Validation Sample			
R-squared	0.068	0.103	0.214
SD relative weights	0.349	0.426	0.626
Sensitivity[b]	0.313	0.375	0.442
Facility-level models of log cost (Validation Sample)[c]			
a. CMI only			
R-squared	0.047	0.136	0.347
CMI coefficient	0.899	1.130	1.290
SE on CMI coefficient	0.125	0.092	0.051
p-value to test coef.=1	0.419	0.155	0.000
b. CMI payment model[d]			
R-squared	0.066	0.153	0.353
CMI coefficient	0.958	1.15	1.29
SE on CMI coefficient	0.124	0.091	0.051
p-value to test coef.=1	0.737	0.095	0.0
c. CMI full model[e]			
R-squared	0.263	0.295	0.396
CMI coefficient	0.627	0.756	1.04
SE on CMI coefficient	0.105	0.082	0.064
p-value to test coef.=1	0.000	0.003	0.486

a. RUG58G contains 8 categories indicating the first digit of the RUG-58 score.
b. % Predicted >=90th percentile given actual >=90th percentile
c. For facility models, cost, CMI, and the area wage index are entered in log form (see methods section in Chapter III).
d. The payment model includes urban/rural status and the area wage index, in addition to the CMI.
e. The full model includes urban/rural status, area wage index, indicators for number of SNF beds (0-9 beds, 10-19 beds, 20-49 beds, 50-99 beds, with 100+ beds as the excluded category), hospital based dummy, ownership type (voluntary/nonprofit dummy, for profit other dummy, and government dummy, with for profit chain as the excluded category) and percent Medicare days, in addition to the CMI.

Person-Level Models

The first panel in Table IV.7 presents statistics of the regression of per diem NTA cost on the variables in each classification system using the test sample data. RUG-58 explains 9.5% of the variance in NTA costs; SIM+RUG58G more than doubles that, to 21.2%. The standard deviation of the relative weights highlights the range of values the 2 models cover. RUG-58 has a standard deviation of 0.429, compared with 0.641 for SIM+RUG58G, indicating that the latter model has a much greater spread in payment values. The sensitivity indicator shows that 37.6% of the highest (10%) cost cases are correctly predicted by RUG-58, compared with 45.4% for SIM+RUG58G. Not surprisingly, both models have higher variance explanation and sensitivity than the RUG-III model.

The second panel in Table IV.7 presents the results for our validation of the models. Comparison of the model statistics based on the test and validation samples indicates that the results are almost identical.

Facility-Level Models

The next three panels present the statistics from facility-level regressions of the log of average per diem costs on the log of the cost-model-based CMI, with (a) no controls, (b) only controls for urban/rural and labor cost differences across areas, and (c) controls for a range of factors typically fixed for a given facility.

The third panel presents the evaluation statistics for the model in which the log of per diem cost is regressed on the log of the CMI, with no controls for differences across facilities.

The RUG-58 explains only 13.6% of the variation in logged facility-level NTA costs; the SIM+RUG58G, in sharp contrast, explains 34.7%, a 2.5fold difference. The CMI coefficient for the RUG-58 model is 1.13, indicating an insignificant increase relative to the desired 1.0 level. The CMI coefficient for the SIM+RUG58G model (1.29) is somewhat higher; and a statistical test rejects the hypothesis that this coefficient equals 1.0 ($p<0.001$). For the RUG-III model, the amount of variance explanation is only about 5%, whereas the CMI coefficient is close to 1.0.

The fourth panel in Table IV.7 presents statistics for the models when conventional Medicare payment adjusters—urban/rural status and wage index—are included with the CMI. Addition of the two variables does not notably affect the evaluation statistics described in the third panel.

The fifth panel presents findings when a number of other facility characteristics are included in the model. These facility characteristics include: (a) hospital-based status, (b) percent Medicare, (c) type of ownership, and (d) indicators of numbers of SNF beds. For the RUG-58 model, controlling for these other factors strongly affects the evaluation statistics. Variance explanation (R-squared), for example, increases to almost 30%, indicating that the facility characteristics added to the model explain a considerable amount of the variation in log of facility-level NTA costs. The coefficient on the CMI drops well below 1.0 and the p-value of 0.003 indicates that it is statistically different than 1.0. With the additional variables, the coefficient on the CMI of the SIM+RUG58G model drops to 1.04, and a statistical test accepts that the coefficient equals 1.0 (p-value=0.486). For the RUG-III model, the variance explanation rises to 26.3%, but the CMI coefficient is only 0.627.

Although not presented in Table IV.7, we conducted a similar evaluation of the CART version of the SIM+RUG58G cost model. Statistics on the CART model mirror those of the regression models presented here. In general, the amount of variance explained by the CART version is slightly smaller than the regression version, with the coefficients on the CMI being generally higher in the CART version.

Models with Charge-based Weights and Case-mix Index

Table IV.8 presents statistics comparable to those presented in Table IV.7, using NTA charges rather than costs to compute the person-level models and relative weights. The facility-level regressions in both Table IV.7 and Table IV.8 are based on a regression of facility-level per diem unadjusted NTA costs on CMIs. In Table IV.8, however, the CMI is calculated from the coefficients in a model of charges, whereas in Table IV.7 it is calculated from costs. We highlight some notable differences between the two tables in the discussion immediately following.

Table IV.8: Comparison of Predictive Models of NTA Charges: RUG-III, RUG-58, SIM+RUG58G[a]

Explanatory Variables	RUG-III (New Weights)	RUG-58	SIM+RUG58G[a]
Person-level models of charges			
a. Test Sample			
R-squared	0.059	0.097	0.283
SD relative weights	0.468	0.598	1.02
Sensitivity[b]	0.300	0.321	0.516
b. Validation Sample			
R-squared	0.063	0.104	0.259
SD relative weights	0.462	0.594	0.993
Sensitivity[b]	0.294	0.320	0.489
Facility-level models of log cost (Validation Sample)[c]			
a. CMI only			
R-squared	0.048	0.144	0.355
CMI coefficient	0.641	0.794	0.866
SE on CMI coefficient	0.090	0.062	0.033
p-value to test coef.=1	0.000	0.001	0.000
b. CMI payment model[d]			
R-squared	0.066	0.161	0.359
CMI coefficient	0.682	0.806	0.859
SE on CMI coefficient	0.089	0.061	0.033
p-value to test coef.=1	0.000	0.002	0.000
c. CMI full model[e]			
R-squared	0.256	0.291	0.396
CMI coefficient	0.369	0.506	0.706
SE on CMI coefficient	0.078	0.058	0.043
p-value to test coef.=1	0.000	0.000	0.000

a. RUG58G contains 8 categories indicating the first digit of the RUG-58 score.

b. % Predicted >=90th percentile given actual >=90th percentile

c. For facility models, cost, CMI, and the area wage index are entered in log form (see methods section in Chapter III).

d. The payment model includes urban/rural status and the area wage index, in addition to the CMI.

e. The full model includes urban/rural status, area wage index, indicators for number of SNF beds (0-9 beds, 10-19 beds, 20-49 beds, 50-99 beds, with 100+ beds as the excluded category), hospital based dummy, ownership type (voluntary/nonprofit dummy, for profit other dummy, and government dummy, with for profit chain as the excluded category) and percent Medicare days, in addition to the CMI.

Person-Level Models

The ranges of predicted charges in Table IV.8, as reflected by the standard deviation of the relative weights, are much wider than those of predicted costs in Table IV.7. The wider ranges were found regardless of whether the comparisons were for the RUG-58, SIM+RUG58G, or RUG-III.

The amount of variance explained (R-squared) at the person level was about the same for the RUG-58 and RUG-III models, regardless of whether the models predicted costs or charges. In the case of SIM+RUG58G, the model based on charges for relative weights explained notably more variance than the model based on costs (0.28 vs. 0.21).

Facility-Level Models

The amount of variance in logged NTA costs explained by each model was about the same whether relative weights were based on costs or charges. The most notable difference between the models in the two tables was that the CMI coefficients in the charge-based models were uniformly less than and statistically different from 1.0, while most of the CMIs in the cost-based models were greater than and only sometimes significantly different from 1.0. In the case of the CMI payment model, for example, the CMI coefficient for the SIM+RUG58G model was 1.29 for the cost-based model and .86 for the charge-based models. These findings suggest that, *if relative weights are based on costs*, high-CMI providers would receive payments lower than costs, while low-CMI providers would receive payments higher than costs. *If relative weights were based on charges*, however, high CMI providers would receive payments higher than costs, while low-CMI providers would receive payments lower than costs. This latter finding seems to reflect different charging practices across SNFs, as indicated by the higher standard deviation of the relative weights in the charge-based models.

E. Additional Findings

In exploring other patient characteristics that are potentially good predictors of NTA costs, we discovered some variables that were statistically significant predictors, but were either exceptionally susceptible to gaming, ambiguously defined, or too cumbersome for a RUG-based approach. These weaknesses notwithstanding, we include the following results to help inform future development of patient assessment instruments and other approaches for creating patient classification systems:

- Number of medications alone yields an R-squared of 10% in a model of log costs; it also adds considerably to models with other regressors (e.g., increases R-squared from 15% to 21%)
- The indicator of radiography on claims has an R-squared of 6% in a model of log costs; its inclusion increases the R-squared for the SIM model of log NTA costs from 18% to 21%
- Inclusion of the indicator of 2+ injections from the MDS leads to an R-squared of 2.6% on its own
- Including a measure of number of physician visits from the MDS leads to an R-squared of 4.5% on its own
- Including disease diagnoses from the MDS leads to an R-squared of 4.6%
- Measures of length of stay or numbers of assessments are negatively related to NTA costs

F. Discussion

The RUG-based approaches represent minimalist options for improving the existing SNF PPS. They address only total NTA costs, essentially leaving the RUG-III intact for nursing and rehabilitation therapy services. The weights from the RUG-based approaches may be considered in concert with additional payment adjustments, such as outlier payments.

Our RUG refinement research led us to present two 'preferred models' for classifying SNF patients according to NTA costs within the RUG refinement approach. The first is the RUG-58, which, as noted, was originally developed by Abt Associates (2000) and is well known by both the policy and provider communities. Additionally, given the classification mechanism currently in use, implementation of RUG-58 would not be perceivably difficult.

The second is the SIM model, which is based on variables used in the RUG-58, changing definitions in a way that doubles the amount of variance in NTA costs explained. This model is based, in large part, on service use indicators defined using the intersection of information from the MDS assessments and the corresponding SNF claims. We initially resorted to this strategy to circumvent the problem that the indication of service use in the MDS assessments could include services received during the prior hospital stay. Our analyses suggest that the intersection of the two does not simply exclude the hospital data, but may also reflect a differential character of cases for which both claims are filed and MDS forms are being completed. One possible explanation is that having both indicators is reflective of service use so substantial that it could not be easily overlooked by either the MDS assessment or the billing process.

Ideally, the service use variables in the SIM would be based on information from the current MDS questionnaire; but the confusion created by the MDS questions presents an important and intractable problem. The substantial improvement seen when restricting the service measures to the relatively small share of such stays with claims indicates that changing the MDS question to capture the circumstances of those indicated by both measures would be worthwhile. The easiest change would be to restrict the relevant service measures to the period within the SNF. In addition, learning more about the circumstances under which both claims and the MDS indicate IV medication or respiratory care in SNFs could provide insight into ways to better measure significant use.

Individuals with indicators for IV medications or respiratory service use tend to have higher NTA costs simply because they incurred costs for IV medications or respiratory services, which tend to be expensive. But it is important to note that use of those two services also increases NTA costs net of the service costs *per se*. In other words, use of IV medications or respiratory services is not only a direct contributor to total NTA costs, but also an indicator of higher than average costs for other NTA services.

The policy-related characteristics of RUG-58 and the current system are essentially the same. Policy-related characteristics of the SIM+RUG58G model include the following:

Statistical Significance

The amount of NTA cost variance explained by the SIM-related models, approximately 20%, is considerably higher than that explained by the WIM-related models developed in prior research. This amount of variance explanation is in the "ball park" with the R-squared derived for case-mix classification systems for Medicare payment for long-term care hospitals and psychiatric hospitals, although it is considerably lower than the R-squared derived for inpatient rehabilitation hospitals.

Clinical Meaningfulness

The SIM variables generally suggest conditions that would lead to high NTA costs. IV medications, for example, are probably more expensive in general than orally administered medications. The respiratory therapy-related variables are another example, given the additional costs for respiratory therapists that are indicated by claims-specified indicators. Although respiratory services may be performed by nurses in some SNFs, it is likely that more complex respiratory therapy services are performed by registered, respiratory therapists.

Administrative Burden

The amount of additional administrative burden associated with these models is relatively small, consistent with their minimalist approach. For the data problem of MDS confounding, information needed to identify NTA classification groups could be developed directly from a revised MDS assessment instrument, minimizing administrative burden.

Clinical and Financial Incentives

The inclusion of service variables in our RUG refinement models inevitably suggests the potential for gaming, whereby providers will have incentives to provide an unnecessary service to increase payment. The TAP for the Abt Associates study helped define the WIM variables in ways that reduce this propensity. Consultant clinicians on the SIM variables expressed mixed opinions about the incentives provided by some of our variables as they are currently defined.

Ease of Implementation and Administration

The SNF PPS has been in effect since 1998, and most nursing homes are familiar with the RUG-III classification system. Since the RUG-III system would remain intact for nursing and rehabilitation therapy with the models presented in this chapter, SNFs would not have to adjust management practices and information systems to manage those service centers.

In sum, either of the two RUG refinement models described in this chapter presents a plausible option for improving the SNF PPS so that it better accounts for NTA costs. By far, the degree of variance explanation identified in the more complex SIM model is superior. But it requires refinements in the MDS assessment instrument to clearly specify that particular services were provided during the SNF stay rather than the prior hospital stay. Clarification of the definitions of types of services used, notably in the area of respiratory care and administration of IV medications and solutions, is also essential for implementation of the SIM model.

V. New Profiles Approach

A. Aims and Background

The goal of the New Profiles (NP) approach is to develop case mix payment methods for NTA and therapy costs in the context of a new classification system based upon clinically meaningful patient groups, defined by patient characteristics. This work, the most ambitious of the five approaches described in the report, is intended to address some of the concerns about the SNF PPS described earlier.

Our analyses had three basic aims. First, we developed an overriding framework for classifying SNF patients into clinically meaningful groups defined by patient characteristics (i.e., the three NP patient classification groups or categories defined in Section B). Second, we developed a patient classification model for NTA costs based on patient characteristics, prior hospital factors, and facility characteristics including the three components of NTA: drugs, respiratory therapy, and other NTAs. Third, we developed a patient classification model for therapy costs (physical, occupational, and speech therapy) using patient characteristics, prior hospital variables, and facility-level variables to determine the potential for a classification system that is not driven by minutes of therapy provided to patients in the SNF (which is potentially gameable).

We did not attempt to develop a patient classification system for nursing costs using the new profiles because extant data sources do not include patient-specific data on such costs. The possibility of combining nursing costs with this approach, however, could be considered in the future.

In our development work, we followed several key precepts:

- The new system must be clinically meaningful to providers
- The new system must not use measures of services endogenous to the facility's ability to control, such as therapy minutes provided, but rather must utilize measures of patient characteristics and conditions. Service measures that are physician driven and not gameable (e.g., IV medication, tracheotomy) are acceptable, excluding those that may be at the facility's discretion (e.g., tube feeding)
- In developing the new system, we should examine separate structures for each cost component (e.g. PT/OT, speech and language pathology, drugs, respiratory therapy) that could then be aggregated as appropriate in patient classification model development

The three major sections of this chapter describe the development of the NP patient classification categories and the development of models relating patient characteristics and other factors to resource use within and across the NP categories as the basis for possible SNF payment system refinements. The models and findings are presented first for NTA and then for therapy services.

B. Development of the New Patient Categories

Conceptual Development

The first step in developing the NP patient classification categories was to define conceptually the major categories of SNF patients. For this purpose, we considered basing some portion of the new approach on classification systems used in rehabilitation hospitals; however, our findings confirmed previous work establishing that separate classification systems are needed for rehabilitation hospitals and for SNFs (Eilertsen, et al. 1998, Buchanan, et al. 2004). The recently developed rehabilitation impairment categories (RICs) do not map effectively to SNF patients. For example, more than 50% of SNF patients were either mapped into the "miscellaneous" RIC or had missing information, precluding placement into any category. Even after consolidating the twenty-one RICs into broader categories that better represent SNF patients, the only two well-defined categories were stroke and hip fracture, and over two-thirds of SNF patients were mapped into the medically complex and unclassified category. Finally, the explanatory power of a CART model using these modified RIC categories was very poor. Thus, consistent with prior work, we concluded that the present day rehabilitation hospital classification system is not directly applicable to SNFs.

Our conceptual development work, therefore, continued with specification of the major categories of patients admitted to SNFs. This categorization was based on input from clinicians about the different types of patients admitted to SNFs from the perspective of the types of care they require, and to a lesser extent the likely duration of their SNF needs and their eventual placement. The conceptual design was not based on information related to resource use, nor did it start with diagnoses. Three broad types of SNF patients were conceptualized:

- The *rehabilitation group* consists of patients admitted primarily for rehabilitative services such as physical or occupational therapy to improve or restore function. Such patients may have diagnoses such as hip fracture or stroke, disability from medical/surgical conditions, and need intensive rehabilitation to regain independent function before being discharged home, or to a more independent environment than a SNF

- The *acute group* consists of patients admitted for skilled nursing care following an acute medical or surgical event such as an MI or abdominal surgery. These patients have primarily skilled nursing care needs (such as wound care or intravenous medications), which are likely to be of relatively limited duration

- The *chronic group* consists of chronically ill or deconditioned patients admitted for skilled nursing care following a hospitalization for a condition such as CHF or COPD, who require assistance with ADLs. These patients often require frequent rehospitalizations and SNF stays or may be long-term nursing facility residents for whom discharge home is not realistic

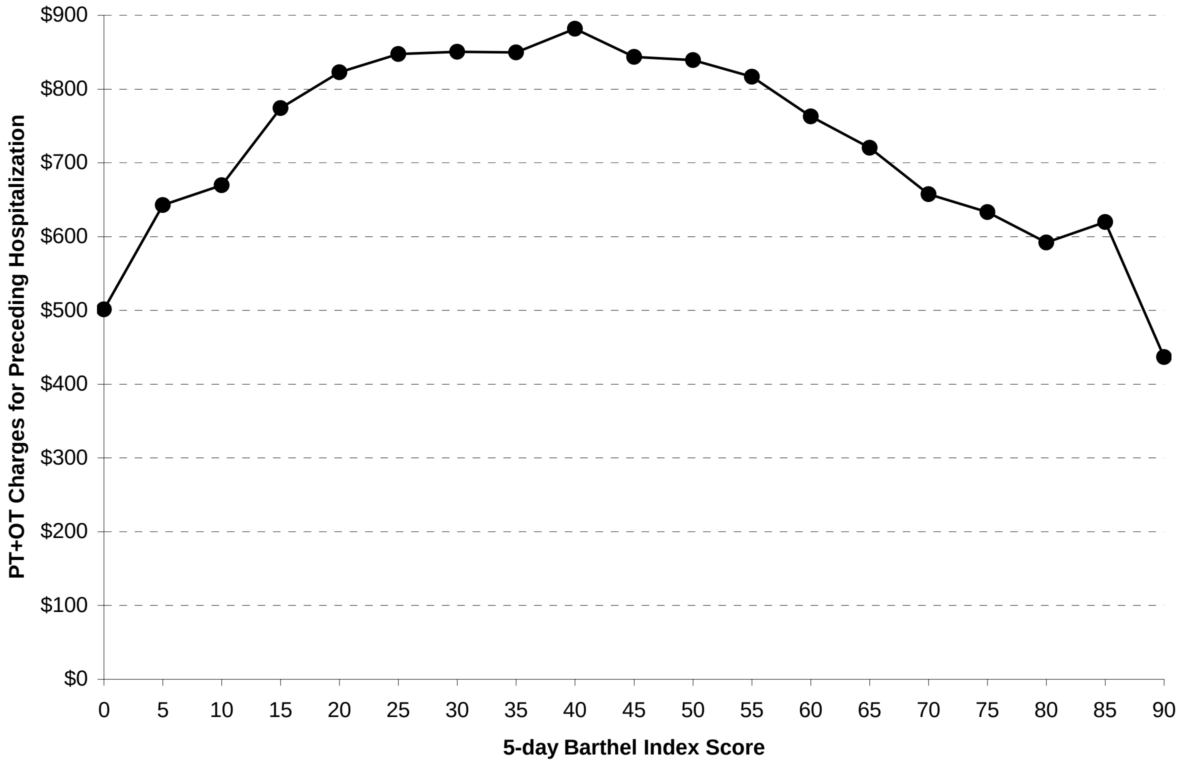
Operationalizing the Group Definitions

There is a continuum of patient conditions and some patients may fit into more than one group. Nevertheless, these groups represent the major types of care commonly provided in SNFs today. Therefore, we went ahead and used data from the preceding qualifying hospital stay and from the first MDS assessment from the SNF stay to operationalize the conceptual definitions of the three groups.

The rehabilitation group was defined based on functional impairment level. To measure functional ability, we used the Barthel Index (Mahoney and Barthel 1965) score calculated from items on the five-day MDS assessment. This is an index of functional ability, which captures several aspects of ADL capability and normally ranges from 0 (worst function) to 100 (best function). Because our functional data were limited to MDS items, our Barthel Index perforce omits an item on the original index (ability to climb stairs), and therefore has a maximum score of 90.

We investigated the relationship between function and therapy services received using the Barthel Index score from the five-day MDS assessment and several different measures of physical and occupational therapy. Physical and occupational therapy were measured (separately and combined) using charges from the SNF stay, derived costs from the SNF stay, and charges from the preceding hospital stay. In all cases, the relationship between function and therapy was shown to be non-linear, wherein patients with moderate functional impairment received more therapy than patients with either substantial or minimal impairment. Figure V.1 is an illustrative example of the relationship between function and the sum of physical and occupational therapy charges for the hospital stay preceding the SNF stay. Patients with a Barthel Index score of 0-10 (extreme impairment) or 70-90 (minimal to no impairment) had substantially lower charges on average ($207 lower for the 0-10 group, $219 lower for the 70-90 group) than patients with a Barthel Index score of 15-65 (serious to moderate impairment). We therefore used a Barthel Index score of 15-65 to define the rehabilitation patient group.

Figure V.1: Relationship between Function and Therapy



Note: Graph is based upon the test sample, prior to application of data exclusion rules

The relationship between function and therapy received in the SNF was very similar to that between function and therapy charges in the prior hospital stay; patients with moderate functional abilities received much greater therapy than patients at either end of the functional scale. Because the acute hospital PPS does not specifically reimburse for therapy services, therapy received in the hospital may be a good indicator of patients' actual need for rehabilitation services.

For patients not included in the rehabilitation group based upon function, our first step in identifying acute conditions was to use the primary diagnosis given on the last claim of the preceding qualifying hospitalization. As already noted, diagnoses listed on hospital discharge records tend to be more accurate and more complete than diagnoses listed on SNF admission records, plausibly due to incentives inherent in the DRG system (or lack thereof in the SNF PPS). Potential acute conditions were identified from ICD-9-CM codes to include diseases in each major diagnostic category that can be an acute event. Only diseases that were clearly chronic in nature (e.g., diabetes mellitus) were excluded from our potential acute diagnosis list. The ICD-9-CM categories shown in Table V.1 were considered potential acute conditions.

Table V.1: Potential Acute Condition ICD-9-CM Codes

ICD-9 Code	Diagnosis Description
001 to 139	Infectious and parasitic diseases
210 to 229	Benign neoplasms
308 to 309	Acute stress and adjustment reactions
320 to 326	Inflammatory diseases of the central nervous system
360 to 389	Disorders of the eye, adnexa, ear, mastoid process
410 to 414	Ischemic heart disease
420 to 459	Cardiovascular disease, cerebrovascular disease
460 to 466	Acute respiratory infection
480 to 490	Pneumonia, influenza, bronchitis
494 to 519	Diseases of the respiratory system
520 to 555	Diseases of the digestive system
557 to 570	Diseases of the digestive system
572 to 579	Diseases of the digestive system
580 to 584	Nephritis, acute renal failure
586 to 629	Diseases of the genitourinary system
630 to 677	Complications of pregnancy, childbirth, puerperium
680 to 706	Infections of skin, subcutaneous tissue
708 to 709	Disorders of skin, subcutaneous tissue
711	Arthropathy associated with infections
730 to 732	Infections of the bone
760 to 779	Conditions originating in the perinatal period
780 to 796	Symptoms, signs, ill-defined conditions
798 to 799	Sudden death, cause unknown
800 to 999	Injury and poisoning
V41 to V49	Conditions influencing health status
V50 to V59	Encountering health services for procedures and aftercare
E codes	External causes of injury and poisoning

Diagnosis alone is not sufficient to distinguish acute from chronic patients, however, because many of the diagnoses on the list of potential acute conditions can also occur in a more chronic form—pneumonia, urinary tract infection, and heart failure are three illustrative examples. In our conceptual framework, the chronic group also includes patients with an underlying problem (or problems) that may require multiple hospitalizations to stabilize recurring acute episodes of what is a chronic disease; COPD is the classic example. Our definition of chronic, therefore, includes not only patients with non-acute diagnoses, but also patients with acute diagnoses who have been hospitalized and admitted to a SNF in the six months prior to the qualifying hospitalization for the current SNF stay. This item is determined by reviewing the entire stream of MDS assessments for each resident, using (a) MDS items AA8a and AA8b to determine if the assessment was a SNF PPS assessment, and (b) A3a, the assessment reference date, to confirm that the assessment occurred during the six-month period.

A flow chart depicting the patient classification scheme is shown in Figure V.2. Using the five-day Barthel Index score, the hospital diagnosis, and prior hospitalization history, we classified patients hierarchically into one of the three broad groups; rehabilitation, acute, or chronic. As noted earlier, patients were assigned to the rehabilitation group if their five-day Barthel Index score was between 15 and 65, accounting for 67% of the total sample. The remaining patients were then split into the acute (17%) and chronic (15%) groups based on their diagnosis and history. The chronic group comprises relatively equal groups of patients with non-acute diagnoses (55% of the chronic group) and patients with acute diagnoses and repeated hospital and SNF admissions (45% of the chronic group). Less than 1% of the sample could not be classified because of missing data.[17]

[17] Additional investigation into the unclassified patients showed that all were missing a 5-day Barthel Index score. Of these, 31% were discharged prior to having a 5-day MDS assessment completed; the remainder had missing or out-of-range values for at least one of the 10 variables used to calculate the Barthel Index score.

Figure V.2: New SNF Patient Classification System

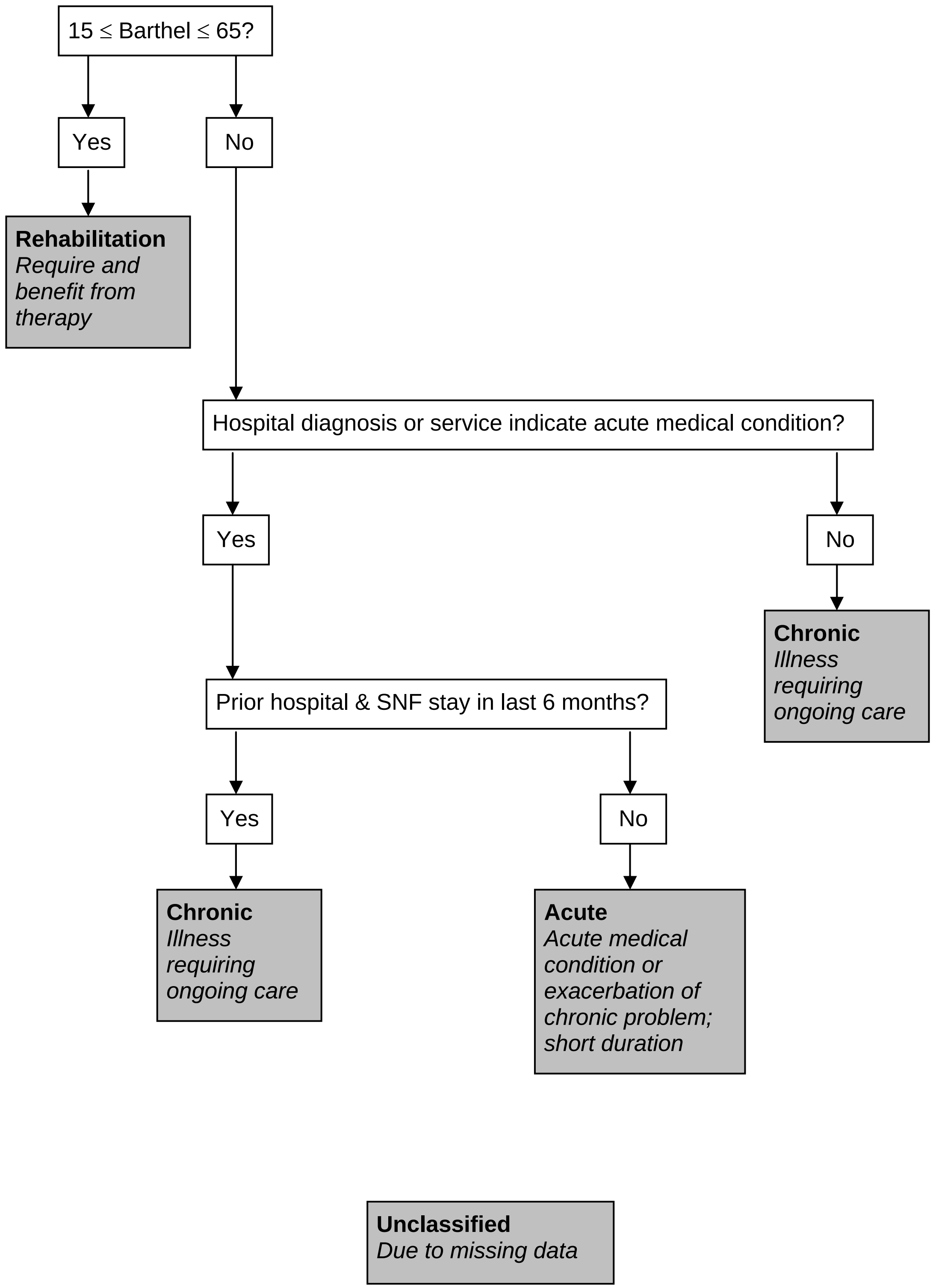


Table V.2 shows characteristics of residents in each of the patient groups. The distribution of diagnoses indicates that even traditional rehabilitation diagnoses (such as hip fracture and stroke) are prevalent in the acute group. This is because such patients may be so frail or injured that they can tolerate only limited rehabilitation services and are receiving mostly nursing care. As expected, medical diagnoses (such as respiratory infection and electrolyte imbalances) occur with greater overall frequency in the acute and chronic group. Dementia and mental disorders are also more prevalent in the acute and chronic group, and skin ulcers are more prevalent among chronic patients, as we expect. Overall, musculoskeletal disease occurs with higher frequency in the rehabilitation group. The Barthel scores on admission are relatively similar between groups, but the rehabilitation group demonstrates increased Barthel scores at fourteen and thirty days, representing greater independence, whereas the acute and chronic groups are more dependent at 14 and 30 days, on average. Thus, function improves more in rehabilitation patients that remain in the facility for 14-30 days than in patients in the other groups. Acute and chronic patients are more likely to have swallowing problems and feeding tubes (particularly chronic patients), and to have severe cognitive impairment. A significant portion of all three groups receive therapy, as we would expect due to the current payment system, but the proportion is higher in the rehabilitation group than in the others. A higher proportion of the rehabilitation group is also discharged home in thirty days, while higher proportions of the other two groups die within thirty days. Thus, diagnoses, function, and selected services and outcomes follow patterns that would be expected given our conceptual classification and group definitions.

Table V.2: Selected Characteristics of the Three Patient Groups in a New SNF Patient Classification System

Characteristic	Rehabilitation Group (n=115,085)	Acute Group (n=29,028)	Chronic Group (n=26,002)
Demographics:			
Average age	80.2	80.3	78.7
Female	67.0	64.1	63.0
White	90.0	85.3	81.8
Married	30.0	28.6	29.2
SNF admission in previous 6 months	22.5	0.0	57.2
Long-term nursing home resident	12.0	25.3	25.1
Diagnoses:			
Acute diagnosis	72.2	100.0	44.6
Hip fracture	10.2	9.4	2.3
Stroke	6.4	15.2	4.2
Respiratory infection	15.0	25.0	19.6
Electrolyte imbalance	27.8	31.6	36.7
Dementia	17.3	27.7	27.6
Mental disorder	31.0	37.7	39.9
Osteoarthritis	14.0	7.7	13.3
Skin ulcer	7.2	8.1	14.9
Musculoskeletal disease	34.3	22.5	29.8
Function and cognition:			
Average 5-day Barthel Index score	39.5	34.5	34.8
Average 14-day Barthel Index score	42.9	30.0	30.3
Average 30-day Barthel Index score	43.0	22.6	24.7
Average % of stay with swallowing problem (K1b)	11.9	29.4	29.7
Average % of stay with feeding tube (K5b)	3.6	16.1	20.4
Average % of stay with Stage 4 ulcer (M1a)	5.7	9.2	17.8
Average 5-day Cognitive Performance Scale	2.5	3.8	3.8
5-day cognition intact	41.3	28.7	29.2
5-day cognition very severely impaired	1.1	19.6	20.2
Therapy received:			
PT or OT received in hospital	77.9	65.1	55.4
PT or OT received in SNF	91.0	75.5	67.9
ST received in hospital	12.8	23.9	15.4
ST received in SNF	17.4	25.9	20.0
Length of stay and 30-day outcomes:			
Average hospital length of stay	8.9	10.2	9.5
Average SNF length of stay	25.2	24.2	22.3
Discharged home within 30 days	33.9	30.3	27.3
Re-hospitalized within 30 days	21.0	20.5	23.5
Deceased within 30 days	8.6	15.9	18.1
Facility characteristics:			
Hospital-based SNF	25.0	21.6	19.7
Urban SNF	78.5	77.5	79.5
For-profit SNF	59.4	63.0	65.5
SNF member of chain	59.3	61.7	60.7
Average SNF total beds	128.2	131.8	138.0
Average SNF % Medicare days	28.6	25.2	24.1

Note: Figures are based upon the test sample, prior to application of data exclusion rules, and are percentages unless otherwise noted.

The average per diem charges and wage-adjusted costs of patients in the three SNF groups are provided in Table V.3. Rehabilitation patients had modestly higher routine charges (reflecting facility-level rates), as well as substantially higher person-level therapy charges, as expected. Also as expected, NTA charges, including drugs and respiratory therapy services, were greater for acute and chronic patients. Wage adjusted costs follow the same pattern.

Table V.3: Average Per Diem Charges and Wage-Adjusted Costs of the Three Patient Groups in a New SNF Patient Classification System

Item	Rehabilitation Group (n=115,085)	Acute Group (n=29,028)	Chronic Group (n=26,002)
SNF stay charges per diem:			
Routine charges	$245.88	$235.41	$233.77
Therapy charges	132.63	99.33	85.41
NTA charges	136.84	149.82	155.85
Total charges	515.34	484.55	475.03
SNF stay wage-adjusted costs per diem:			
Routine costs	204.51	194.15	191.04
Therapy costs	77.71	58.08	50.76
NTA costs	57.43	63.67	70.15
Total costs	337.07	314.90	310.18

Note: Statistics are based upon the test sample, prior to application of data exclusion rules

Discussion

A key feature of this new classification system is that patients are assigned to the three groups based solely on their clinical and functional characteristics and their medical history. We explicitly avoided including measures of service provision, such as tube feeding or physical therapy, since inclusion of such items can lead to tautological relationships of dependent variables (e.g., using the number of therapy minutes provided to predict the amount of therapy resources used) or create perverse incentives to provide sub-optimal care (e.g., encouraging the use of urinary catheters by providing higher reimbursement for such patients).

In this data set, the missing data were confined to the five-day Barthel Index scores, either because patients were discharged prior to completion of a five-day MDS assessment, or because there were problematic data on the assessment. In practice, however, missing data could also arise from either of the other two items used in the classification scheme—hospital diagnosis or SNF admission history. In our analysis, such stays (7.5% of all stays) with missing data were eliminated from the data set. Our analysis had the advantage of a data file containing pre-matched qualifying hospitalization claims and SNF stays, as well as the entire stream of MDS assessments for each resident. Such information will not be known with the same level of accuracy by SNF admission coordinators or assessment nurses; however, implying that actual implementation of this classification system may require a different approach.

From a conceptual perspective and in terms of clinical profiles of the groups, the proposed patient classification is reasonable. The differences in charges and costs for different components of SNF care reflect expectations. The next step is to use these groups as the basis for models explaining variance in SNF stay-level costs and charges per day for NTA and therapy services.

C. NP-NTA Models

One goal of a revision to the current SNF PPS is increasing the ability to accurately pay for NTA costs. The RUG-III classification system has been shown to explain only about 5% of the variance in NTA resource use (Kramer, et al. 1999, Abt Associates 2000, Liu, et al. 2002, Fries 2002). The impact of the disparity between payment and actual resource use includes SNF reluctance to accept patients with large anticipated NTA costs. A recent OIG report cited expensive drugs and IV antibiotics as the most common reasons discharge planners were unable to place a patient in a SNF. Chemotherapy and HIV patients were also cited as difficult to place; discharge planners alleging the high cost of drugs as a contributing factor (OIG 2001). Indeed, drug costs are the greatest contributor to NTA costs. Overall, nursing home residents receive an average of 7.3 routine medications (Briesacher, et al. 2002) and the percentage of residents with at least nine prescriptions rose from 18% to 27% between 1997 and 2000 (Mendelson, et al. 2002).

In this section, we examine the relationship between NTA resource use and an extensive set of potential predictors within each NP group. Our initial purpose is to account for as much variation as possible in resource use for each NP group, and we therefore include variables at this stage that may not be appropriate in a final patient classification system for payment purposes. The list of predictors goes considerably beyond the MDS items considered by Abt Associates (2000) in prior efforts to improve patient classification for NTA services. We present our findings on the determinants of NTA resource use from two approaches: a two-stage regression model and a regression tree (CART) model. We then adapt and revise the models, simplifying them into a model that could form the basis of a payment system.

Data

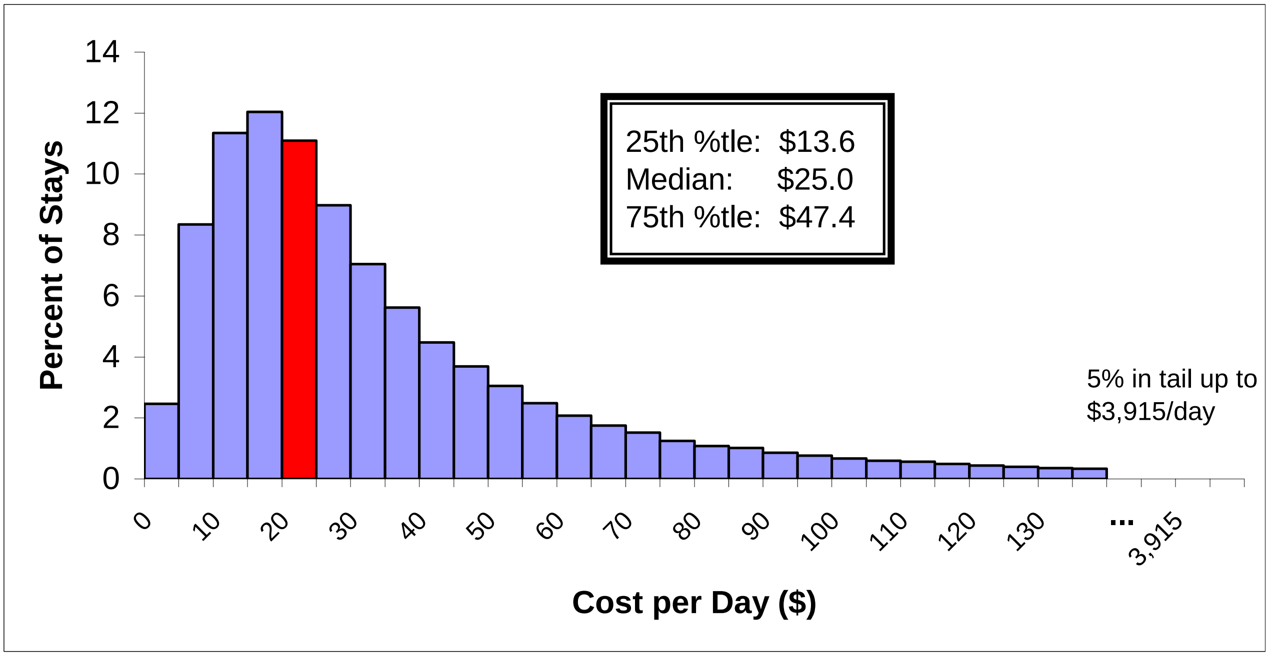
NTA Resource Use Variables

The data for the initial analysis on NTA resource use is the test sample, described in detail in Chapter III. Resource use was measured by per-day charge and cost variables for the three main NTA components: drugs, respiratory therapy and other non-therapy ancillaries (ONTAs). The per-day average of total NTA charges was $141.53, compared with costs estimated at $61.71. Drugs were by far the largest component, with a per-day average charge of $82.63 and cost of $41.44.

All three components had a significant percentage of stays with no charges recorded: drugs (6.4%), ONTA (20.2%), and respiratory therapy (89.0%). The zero charge (cost) stays influenced our statistical model choices. These stays may be zeros either because (a) the patient did not use ancillary resources or (b) the SNF incorrectly assigned ancillary charges in the submitted claim. Our calculations include zeros by treating them as stays for which no resources were used.

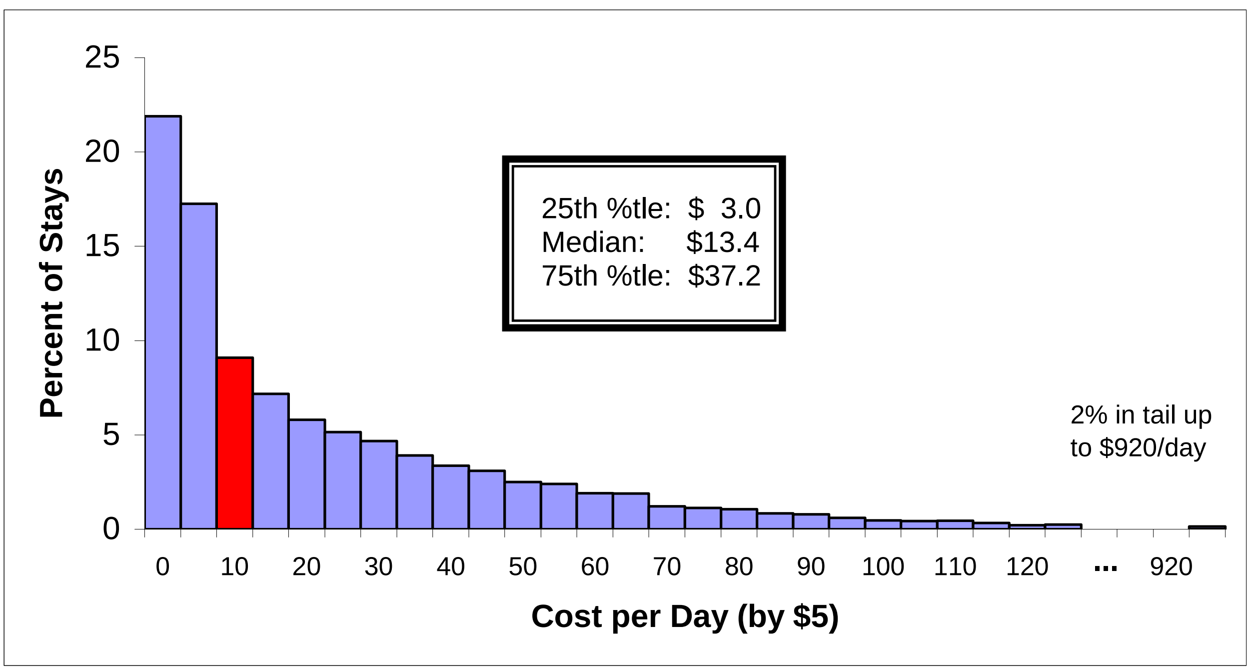
The distributions of costs for the sample stays with positive costs can be seen in Figures V.3 through V.5. These distributions display the enormous cost variation in each of these components. The charge distributions have a similar shape, but a greater range of values. Both charges and costs are also skewed, with broad upper tails indicating that each of the three components had a small percentage of stays that were extremely costly, as can be seen in Table V.4.

Figure V.3: Drug Cost Distribution (All Stays with Costs > $0)



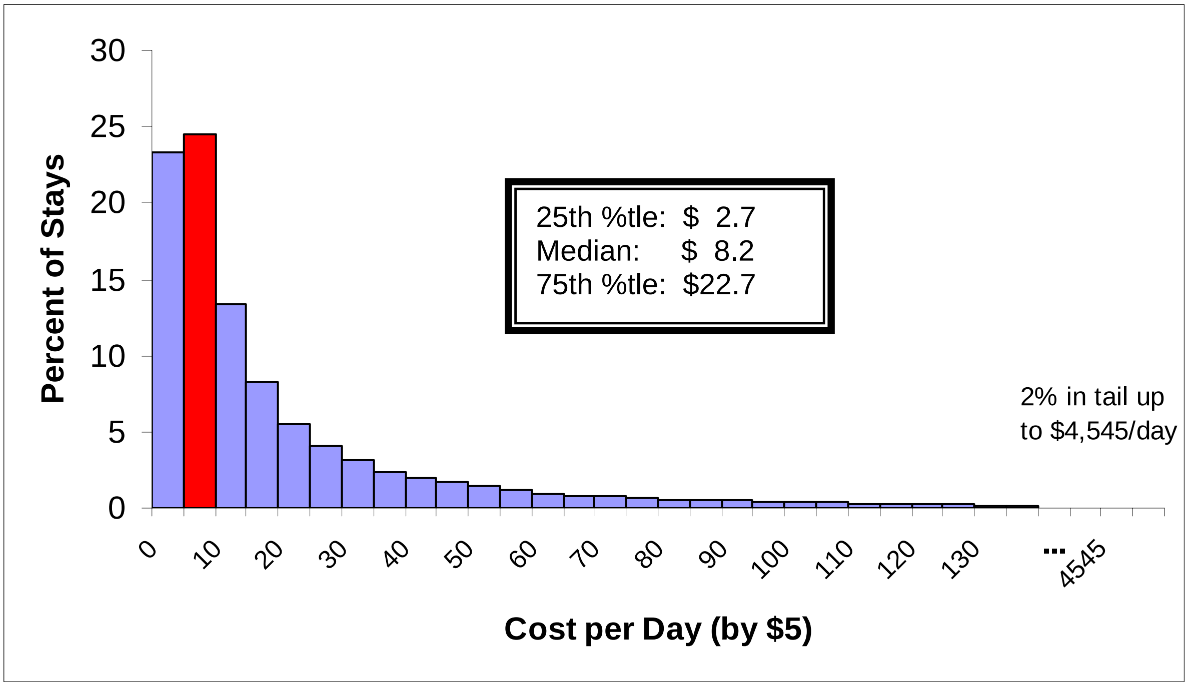
Note: The distribution is based upon the test sample

Figure V.4: Respiratory Cost Distribution (All Stays with Costs > $0)



Note: The distribution is based upon the test sample

Figure V.5: ONTA Cost Distribution (All Stays with Costs > $0)



Note: The distribution is based upon the test sample

Table V.4: Distribution Skew for NTA Costs and Charges

	% w/ no use	95%tile	97%tile	99%tile
Costs				
Drugs	6.4%	$141	$188	$330
Respiratory	89.0%	$91	$112	$220
ONTA	29.2%	$89	$117	$191
Charges				
Drugs	6.4%	$332	$462	$806
Respiratory	89.0%	$376	$465	$716
ONTA	29.2%	$283	$394	$688

Note: Statistics are based upon the test sample

Some potentially important cost factors were evident from comparing high-cost to low-cost outliers. The high-cost outliers differed from the low-cost outliers consistently across components. The high-cost stays had much longer qualifying hospital stays and were more likely to have spent time in the hospital ICU. They also had much shorter SNF stays than the low-cost stays and were much more likely to be re-hospitalized or to have died by thirty days post-admission. Overall, NTA stays with high per day costs appeared to be short stays with severely ill patients that either returned to acute care or died. Those stays with low per diem costs, in contrast, were long stay patients with extensive rehabilitation therapy and more likely to be discharged home.

Although we examined both costs and charges, we focused our initial modeling on charges, since the variance explanation was generally higher when charges rather than costs was the dependent variable. Also, although the skew in the charge distributions suggested the possibility of using logged charges as the dependent variable, we found that the logged values did not significantly improve the fit of our models.

Explanatory Variables

NP patient classification system

The NP patient classification system, described in Section B, assigns patients to one of three primary categories: rehabilitation, acute, and chronic. In these analyses, we first examined resource use separately for each of the three categories. Then, we analyzed resource use for all stays (i.e., combining the three NP groups), treating the categories as explanatory variables in the model.

Other explanatory variables

The explanatory variables considered for this analysis fall into five groups: [18]

- Demographic variables, including age and gender
- Clinical diagnoses, identified through hospital and SNF claims diagnosis codes and from the MDS. These conditions are likely to be associated with specific drug regimens, treatments, and with the need and extent of respiratory therapy and other NTA use. The multiple sources of information generated overlapping diagnosis variables. For example, stroke may be coded in one of several MDC categories, and is also coded from specific ICD-9-CM codes for stroke. Thus, MDC categories, such as MDC 1: Nervous System, overlap with the variable identifying a stroke diagnosis through ICD-9-CM codes for stroke
- Service indicators created from both hospital and SNF claims, including such items as the use of radiology services and the drug charges for the qualifying hospital stay. Discharge from the qualifying hospital stay may be up to 30 days prior to the SNF admission but is generally the most recent stay prior to SNF admission. To the extent that these services are subsequently provided to the patient in the SNF, or identify more complex patients, qualifying hospital service information may predict NTA resource use
- Functional indicators, derived from patient scores on the ADLs. These indicators, for such items as mobility, may indicate NTA resource use of items such as pressure relieving beds and wheelchairs
- Facility characteristics. Some of the differential in costs of care among SNFs may not be fully accounted for by the available case-mix measures.

[18] Descriptive statistics for all the explanatory variables used in our initial NTA analyses are in Appendix V.1 of Urban Institute (2004).

Methods

Two-Stage Regression Models

A two-stage model was used to identify NTA resource relationships with explanatory variables. The first stage identified which patients used any resources; the second stage identified variables associated with the extent of use. Our preliminary analyses demonstrated that the variables associated with these two relationships were different, supporting the two-stage modeling approach.

Variable reduction occurred through four steps:

- All potential covariates were assigned an initial grouping. The groupings created blocks of related variables such as Major Diagnostic Categories. Regression analysis on both the likelihood of any charges and on the charges for those stays with charges greater than zero was performed on each block and for each of the three NTA components
- Selection of variables that were significantly related to charges at a 5% level for multiple components reduced the number of variables from about 600 to 150. This set of variables was then entered into linear probability models for any charges and ordinary least squares models for charge amounts (if greater than zero) using a stepwise selection method to maximize R-squared. Linear probability models are not the best choice for prediction, but they will correctly identify significant relationships between the likelihood of charges and explanatory variables
- Regression models were run separately by NTA component for each patient classification category and for all stays. Variables that increased the adjusted R-squared by at least 0.005 were retained in the final model
- Finally, certain variables that did not increase the R-squared by 0.005 but that had large and significant effects on charges were put back into the model. For example, the indicator for HIV has a large effect on NTA resource use. Since it affects very few stays, it does not contribute much to the explanatory power of the model over all stays. However, excluding it would greatly reduce the model's ability to explain resource use in facilities that treated any HIV patients

This process generated a separate set of predictor variables for each NTA component/patient classification combination and allowed the variables that were most predictive of the likelihood of any charges to differ from those that predicted charge amounts.

These sets of variables were then entered into Heckman two-stage models (Greene 1997). In the first stage, a probit model was used to predict the likelihood of any charges, generating an adjustment factor for each stay. Ordinary least squares were then used to predict the value of the charges, incorporating the adjustment factor. The models were then replicated using cost per day as the measure of resource use rather than charges.

Regression Tree (CART) Models

As discussed in Chapter III, traditional regression models may miss important interactions among explanatory variables as well as relevant non-parametric relationships between resource use and explanatory variables. While clinical expertise hypothesized some of these relationships and we were able to incorporate them into the regression models, we believed some interactions or non-parametric relationships would still be missed. In consequence, CART was applied to the data and used in this study. (See Chapter III for a full description of CART methodology.)

Due to computing limitations, CART models were created using only a 1% random sample of stays. Fortunately, preliminary analysis found the 1% file to retain the general distributions of charges (costs) and explanatory variables from the test sample. (The CART models used the variable set identified in step 2, described above.)

Since the CART method will eventually create as many groups as observations in the data, the trees must be limited through a consistent rule to obtain useful results. The usual pruning criteria of '1 standard error rule' (Breiman, et al. 1984) generally resulted in trees with few nodes. Where possible, trees with up to ten terminal nodes were obtained to more deeply explore the structure. Each terminal node was required to have at least fifty stays. We integrate the findings from the CART and regression models into one approach below by incorporating the nonlinearities and interactions identified by CART into the regression models.

Results of the Initial Models

Findings on the overall fit (explanatory power) of the regression models are presented first. Then, the main variables associated with the amount of NTA charges for each of the three NTA components (drug, respiratory, and ONTA) are covered, followed by the variables associated with any NTA charges. We follow this by noting variables identified by the CART analyses as related to NTA charges. For the cost models, only the findings on model fit are presented, since the cost models contained the same variables, used the same econometric and CART methodology, and yielded the same pattern of results as the charge models.

Model Fit

Table V.5 shows the amount of variation in NTA resource use explained by each model for each NP group and the entire sample. Overall, the two-stage models explained the variations in NTA charges reasonably well, particularly for ONTAs and respiratory therapy. The drug and ONTA results showed similar patterns. The models were able to explain the largest amount of variation in acute stays and the least amount of variation in chronic stays. The cost models generally had poorer variance explanation than did the charge models. The CART models explained somewhat less variation than the regression models, and did not have the same pattern in explanatory power across patient classifications or NTA components as the regression models.

Table V.5: Explanatory Power for All Models

Variable	R-squared Acute (N=27,481)	Chronic (N=24,276)	Rehabilitation (N=111,043)	All (N=162,800)
Drug				
Two-Stage Model				
Charges	0.39	0.29	0.32	0.32
Costs	0.17	0.10	0.13	0.12
CART Model				
Charges	0.21	0.20	0.22	0.21
Costs	0.11	0.09	0.10	0.09
Respiratory Therapy				
Two-Stage Model				
Charges	0.45	0.44	0.51	0.46
Costs	0.47	0.48	0.45	0.48
CART Model				
Charges	0.28	0.15	0.36	0.28
Costs	0.20	0.20	0.21	0.21
Other Non-Therapy Ancillaries (ONTAs)				
Two-Stage Model				
Charges	0.46	0.40	0.42	0.44
Costs	0.31	0.15	0.26	0.25
CART Model				
Charges	0.32	0.35	0.35	0.35
Costs	0.16	0.15	0.17	0.16

Notes: R-squared represents the amount of variation in resource use across all stays explained by the model. Models run on the test sample with excluded observations as described in Chapter III.

Variables Associated with NTA Charges in the Two-Stage Models

Table V.6 shows the average effect on drug charges of variables that contributed, on average, at least $20 of drug charges per day.[19] The effects were on stays that had positive drug charges. For continuous variables, the contribution represents the effect of having the average value of the variable as compared to having a zero value for the variable. One variable was

[19] The complete results of the two-stage Heckman model and the parameter estimates for all variables that appear in each stage of the final models are presented for all patient classification categories, found in Appendix V.4 of Urban Institute 2004.

strongly associated with the likelihood of a stay having positive drug charges, but did not contribute at least $20 of drug charges for stays with drug charges: an MDC of musculoskeletal system. The variables include clinical conditions such as HIV as well as services and functional status variables. They also include only one facility characteristic, hospital-based, and one variable from the qualifying hospital stay, drug charges per day. It also depicts the patient classification categories in which the variable was present.

Table V.6: Variables Associated with at Least $20 of Drug Charges Per Day

Variable	Average Effect on Drug Charges /Day	Models In Which Variable Appears
MDC: HIV	179.74	A,C,R,T
IV Therapy or Solution Use in SNF	147.45	A,C,R,T
Solid Organ Transplant	126.82	A,R,T
Hospital-based Facility	108.53	A,C,R,T
Parenteral IV	55.36	A,C,R,T
Bedfast	37.41	A,C,R,T
Drug charges in Hospital	36.10	A,C,R,T
ADL: Walk in Corridor	33.77	C,R,T
Use of Respiratory Therapy in SNF	32.75	A,R,T
Age >=85	31.21	A,C,T
Suction	29.00	A,C,T
Chemotherapy	25.84	R,T
Foot Infection	23.61	C,R,T
Variables strongly associated with the likelihood of any charges but not in the above list	*Average Effect on the Likelihood of Drug Charges*	
MDC: Musculoskeletal System	0.13	R,T

Notes: Only variables appearing in at least 2 patient classification categories are included in this table. The Average Effect is the size of the effect on resource use per day of the presence of the indicator, or, in the case of continuous variables, the size of the effect on resource use per day of the average per stay value of the variable. The following abbreviations are used for the patient classification models: A=Acute, C=Chronic, R=Rehabilitation, T=Total patients. Models run on the test sample with excluded observations as described in Chapter III.

For respiratory and ONTA charges, the Tables V.7 and V.8 present the variables associated with at least $15 of charges per day, due to the smaller overall charges per day for the respiratory and ONTA components.

Table V.7: Variables Associated with at Least $15 of Respiratory Charges Per Day

Variable	Average Effect on Respiratory Charges /Day	Models In Which Variable Appears
MDC: HIV	195.97	C,T
Ventilator Use	149.78	C,T
Tracheostomy Care	141.86	A,C,R,T
Hospital-based Facility	95.49	A,C,R,T
Adverse Effects of Medication	68.27	A,T
Chronic Obstructive Pulmonary Disease (COPD)	38.95	A,R
Oxygen Use	32.27	R,T
Bedfast	30.41	A,R,T
Respiratory Therapy Use in Hospital	24.21	A,C,R,T
Shortness of Breath	21.34	A,C,R,T
IV Therapy or Solution Use in SNF	19.19	A,R,T
Respiratory Infection	19.05	A,R,T
Variables strongly associated with the likelihood of any charges but not in the above list	*Average Effect on the Likelihood of Respiratory Charges*	
Tracheostomy Care	0.53	A,C,T
Use of CCU in Hospital	0.10	C,R,T
MDC: Respiratory System	0.23	A,C,R,T

Notes: Only variables appearing in at least 2 patient classification categories are included in this table. The Average Effect is the size of the effect on resource use per day of the presence of the indicator, or, in the case of continuous variables, the size of the effect on resource use per day of the average per stay value of the variable. The following abbreviations are used for the patient classification models: A=Acute, C=Chronic, R=Rehabilitation, T=Total patients. Models run on the test sample with excluded observations as described in Chapter III.

Similar to the drug charge variables, the variables associated with large incremental respiratory and ONTA charges were a combination of clinical conditions, services, qualifying hospital information and the facility characteristic, hospital-based. The variables overlapped by NTA component. Three variables—IV therapy or solution use in SNF, bedfast, and hospital-based facility—occurred in all NTA components. Each NTA component had unique contributing variables as well.

Table V.8: Variables Associated with at Least $15 of ONTA Charges Per Day

Variable	Average Effect on ONTA Charges /Day	Models In Which Variable Appears
Hospital-based Facility	167.46	A,C,R,T
IV Therapy or Solution Use in SNF	83.47	A,C,R,T
Use of Respiratory Therapy in SNF	74.35	C,R,T
Dialysis	30.68	A,R,T
Radiation	30.24	C,R,T
Suction	22.92	A,C,R,T
Lab Charges in Hospital	20.68	A,C,R,T
Bedfast	19.44	A,C,R,T
Parenteral IV	19.25	A,C,T
Oxygen Use	18.78	A,C,R,T
Transfusion	16.12	A,R,T
Skin Ulcer	15.10	C,R,T
Variables strongly associated with the likelihood of any charges but not in the above list	*Average Effect on the Likelihood of ONTA Charges*	
Feeding Tube	0.32	A,C,R,T
Application of Dressings	0.20	A,C,R,T
Hip Fracture	0.13	A,R
Deyo Co-Morbidity Index	0.12	A,C,R
Application of Medications	0.12	C,R,T
Surgical Wound Care	0.11	C,R,T
Relief Bed Use	0.10	A,C,R,T

Notes: Only variables appearing in at least 2 patient classification categories are included in this table. The Average Effect is the size of the effect on resource use per day of the presence of the indicator, or, in the case of continuous variables, the size of the effect on resource use per day of the average per stay value of the variable. The following abbreviations are used for the patient classification models: A=Acute, C=Chronic, R=Rehabilitation, T=Total patients. Models run on the test sample with excluded observations as described in Chapter III.

The two-stage model was most effective in explaining the variation in respiratory therapy and ONTA use and charges. The first stage was able to explain nearly 25% and 10%, respectively, of the presence of any use for respiratory therapy and ONTA, but only about 5% for drug use. While many of the variables associated with an increased likelihood of any charges in the respiratory and ONTA models were also associated with large incremental charges per day, some variables only emerged in the first stage of the model, pertaining to the likelihood of any use. Variables that had large effects on the likelihood of any respiratory or ONTA use but not on the size of the charges are presented at the bottom of Tables V.7 and V.8, respectively.

Nor were variables identical across NP patient classification categories (acute, chronic, or rehabilitation) within an NTA component. The tables above show that even variables with large effects on resource use may not have an effect in every patient classification category. Even for those variables that appeared in multiple patient classifications, the size of the effect on resource use varied across classifications. For example, cellulitis has a much larger effect on drug charges per day for chronic stays ($27.07) than for rehabilitation stays ($12.14).

Variables Related to Resource Use in CART Models[20]

The CART models reinforced many of the two-stage model findings and also yielded additional variables created as subcategories or interactions by the regression tree method. These variables were included in the next model development step, moving the models toward a payment system.

Moving the Models toward a Payment System

Using the two-stage model approach, we were able to explain fully one-third to one-half of the variation in charges for the three NTA components at a stay level. Different explanatory variables were significant for each NTA component, and the effect (coefficient) of many of the

[20] Full details of the CART models of NTA resource use can be found in Chapter V.C and Appendix V.5 of Urban Institute 2004.

variables in each NTA component model differed across patient classification categories. Further, variables that were important in determining whether an NTA component was received typically differed from those that were important in determining the per-diem charge (once the service was received). Many modifications are required to move from these models to a model upon which a payment system can be based.

As an initial step, we made the following refinements to the models developed to this point: (1) removed variables considered undesirable in a payment system; (2) consolidated the component models into models for total NTA resource use, converted to single stage models, and incorporated information from the CART models and the first stage of the two-stage models as additional variables; and (3) refined the variable selection process to limit the number of variables in the model. Each is discussed briefly in turn:

Inappropriate Payment System Variables

Since a good payment system must take into account the incentives created by the model, we dropped variables from the models that were inappropriate for a payment system. Examples of these variables can be found in Table V.9, along with the reason for their exclusion.

Table V.9: Variables Excluded from the Model for Payment Related Reasons

Variable	Reason for Exclusion	Attempts to Proxy
Bedfast	This variable creates an undesirable incentive for the SNF not to get patients out of bed	Examined ADL variables for mobility, and combinations of these variables as substitutes
Hospital Based	Unable to use this variable in a payment system for equity reasons	Examined substituting clinical characteristics and interactions found to be associated with hospital-based SNF patients
Malnutrition on MDS	This variable creates an undesirable incentive for the SNF	Substituted variables such as a hospital diagnosis of malnutrition, ADL Eating dependence >= 3, and MDS – Swallowing problems
Feeding Tube	This variable creates an undesirable incentive for the SNF to use feeding tubes	Substituted variables such as a hospital diagnosis of malnutrition, ADL Eating dependence >= 3, and MDS – Swallowing problems

We attempted to proxy for those variables that had large effects on resource use through various mechanisms. For the bedfast variable, for example, we looked at combinations of mobility ADLs. For other variables, such as malnutrition, we were able to substitute diagnosis codes for the qualifying hospital stay, which by definition eliminated the SNF incentive problem.

Consolidating the NTA Components into a Single Model

The first attempt to consolidate NTA components into one model was simply the union of all of the second stage variables contained in any of the three component models, using the substitutions for undesirable variables described in Table V.9. We then added interaction terms derived from the CART model terminal nodes. The two-stage models were originally used because of the concern that respiratory therapy is used in only a small percentage of stays (11%) but when it is delivered, it is often costly (see Figure V.4). Therefore, we added variables to each of the NP category models derived from the interaction of variables in the first stage of the respiratory therapy models that increased the likelihood of a patient receiving respiratory therapy by 10%. (These variables can be found in Appendix V.4 of Urban Institute 2004.) Interacting each of six variables with all the others created up to fifteen potential variables that represented the presence of two conditions or services—such as a variable indicating shortness of breath in conjunction with a diagnosis of COPD. We also included a variable suggested from the SIM specifications: the interaction of respiratory therapy in SNF claims and oxygen use in the MDS.

Refining the Variable Selection Process

The resulting models contained far too many variables for an efficient payment system and we were forced to significantly reduce the number, with a goal of less than 30 variables in each final model. We eliminated variables using several approaches:

- A statistical selection process generated the order of addition to the models for each variable in order to maximize the fit (R-squared) of the models. We were able to drop many variables that contributed little to the fit of the model. Some variables such as organ transplant status were retained in the model, even though they had a small impact on overall fit, because they were associated with extremely high costs although found in only a few patients
- Variables that entered the model in an unexpected and undesirable direction were dropped. For example, the use of a relief chair was associated with lower NTA costs in rehabilitation patients. Since we do not want to discourage the use of relief chairs, the variable was dropped from the model
- Variables that were too complex for a payment system based on adjustment factors were also dropped. These were defined as interaction terms identified by the CART models that involved the combination of more than three variables
- After variable elimination, we converted continuous variables from the qualifying hospital stay into categories. For example, hospital drug charges were divided into five categories based on the percentile in which the drug charges fell: 0-50th, 51-75th, 76-90th, 91-95th, 96-100th. This ensures that each SNF stay falls into only one of the five categories. It has the added advantage of reducing some of the potential incentives for hospitals to change their drug charging process in order to make their discharges more appealing to SNFs.

Final Models

Variables remaining in the final stay-level models are listed in Table V.10. Means are provided for scales and continuous variables; percentages are provided for dichotomous variables. The final stay level model for each of the three NP groups is presented in Table V.11. Ordinary least squares models were run using both wage-adjusted costs and wage-adjusted charges as dependent variables, and the coefficients from the cost models are presented in the table. They represent the average effect on the per diem non-therapy ancillary cost. For example, a solid organ transplant patient in the acute profile group is, on average, $70.49 more costly per diem than a patient without an organ transplant—not surprisingly, since these patients need expensive drugs.

Table V.10: Means for Variables in the Final NTA Models

Variables	Mean (% unless indicated)
Qualifying Hospital Variables	
Diagnosis - Chronic pulmonary disease	24.23%
Diagnosis - Hip fracture or osteoarthritis	20.99
Diagnosis - Malnutrition	6.14
Diagnosis - Renal failure	10.54
Diagnosis - Sepsis	5.34
Diagnosis - Skin ulcer	8.41
Diagnosis - Solid organ transplant	.13
Diagnosis - Cellulitis	4.92
MDC - HIV	.10
MDC - Infection but not sepsis	11.32
MDC - Myeloproliferative	9.67
Hospital per day drug charges >= 95th %tile	4.97
Hospital per day drug charges 90-95th %tile	4.99
Hospital per day drug charges 75-90th %tile	24.94
Hospital per day drug charges 50-75th %tile	25.02
MDS Variables	
Chemotherapy	.93
Dialysis	2.88
Foot infection	1.91
# of days with injections	1.59 days
Oxygen	31.49
Parenteral IV	7.42
Radiation	.76
Suction	2.46
Tracheostomy care	1.03
Shortness of breath	19.47
Transfusion	5.91
Vomiting	3.55
# of mobility ADLs >= 3	3.53 ADLs
ADL Scale - Bed mobility (0-4)	2.02 units
ADL Scale - Eating (0-4)	1.14 units
ADL Scale - Eating >= 3	22.74
ADL Scale - Locomotion off unit (0-4)	2.81 units
Barthel score (0-90)	38.16 units
Interaction Variables	
Use of IV in claims and on the MDS	9.10
Use of respiratory therapy in claims and MDS oxygen use	6.82
MDS Injections > 1 and no IV use in claims	31.75
No use of IV in claims, ADL - Dressing < 4 and no MDS - Application of dressings	72.38
Use of IV in claims and Hospital diagnosis - COPD	4.72
Use of IV in claims and Hospital lab charges per day > $499	2.26
Use of IV in claims and Hospital MDC - respiratory system	5.32
Use of IV in claims and MDS - Relief bed use	3.22
MDS - Shortness of breath and MDS - Oxygen	15.01
MDS - Oxygen and Hospital MDC - respiratory system	20.87
Hospital Diagnosis - COPD and Hospital MDC - respiratory system	23.42

Note: Statistics are based upon the test sample, and are percentages unless otherwise indicated.

Table V.11: Final NTA Models: Variables and Their Dollar Effects on Cost Per Day

Variables	Acute	Chronic	Rehabilitation
Qualifying Hospital Variables			
Diagnosis - Chronic pulmonary disease	7.66		7.26
Diagnosis - Hip fracture or osteoarthritis			-5.04
Diagnosis - Malnutrition		5.62	
Diagnosis - Renal failure		5.19	
Diagnosis - Sepsis			8.76
Diagnosis - Skin ulcer	6.51	5.38	11.34
Diagnosis - Solid organ transplant	70.49		79.14
Diagnosis - Cellulitis	10.51		
MDC - HIV	153.78	138.79	101.67
MDC - Infection but not sepsis	11.19		7.15
MDC - Myeloproliferative	11.45		
Hospital per day drug charges >= 95th %tile	36.19	50.75	39.43
Hospital per day drug charges 90-95th %tile	32.14	36.48	27.73
Hospital per day drug charges 75-90th %tile	21.62	16.75	18.11
Hospital per day drug charges 50-75th %tile	10.39	11.41	10.93
MDS Variables			
Chemotherapy			16.25
Dialysis	20.99		
Foot infection		15.60	15.13
Injections	1.87	2.55	1.52
Oxygen	6.11		
Parenteral IV	31.12	28.03	33.36
Radiation		24.64	
Suction	22.5	18.35	15.88
Tracheostomy care	24.39	35.56	15.11
Shortness of breath	12.02	11.70	
Transfusion	8.69		12.13
Vomiting			17.10
# of mobility ADLs >= 3	4.71		
ADL Scale - Bed mobility	4.51		1.62
ADL Scale - Eating			4.61
ADL Scale - Eating >= 3	7.31		
ADL Scale - Locomotion off unit			3.06
Barthel score	0.68		0.276
Interaction Variables			
Use of IV in claims and on the MDS	68.33	38.80	65.36
Use of respiratory therapy in claims and MDS oxygen use	60.19	62.26	54.39
MDS Injections > 1 and no IV use in claims	8.92		10.59
No use of IV in claims, ADL - Dressing < 4 and no MDS - Application of dressings		-21.96	
Use of IV in claims and Hospital diagnosis - COPD			14.09
Use of IV in claims and Hospital lab charges per day > $499	8.86		16.46
Use of IV in claims and Hospital MDC - respiratory system	21.27	9.38	
Use of IV in claims and MDS - Relief bed use		55.74	
MDS - Shortness of breath and MDS - Oxygen			12.59
MDS - Oxygen and Hospital MDC - respiratory system		8.73	4.17
Hospital Diagnosis - COPD and Hospital MDC - respiratory system		5.20	

Notes: Coefficients are from cost models run on the test sample. All coefficients were significant at P<=.05.

The variables in Table V.11 have been grouped into three categories: (1) variables whose values are determined from the qualifying hospital stay; (2) variables whose values are determined from the MDS; and (3) variables that are a combination of two or more factors.

The variables with the largest effects on NTA resource use are similar to those found in the two-stage models reported earlier in this chapter (see Table V.11). For example, the use of IV increased costs directly by $68.33, $38.80 and $65.36 in the acute, chronic and rehabilitation profile groups, respectively. An HIV diagnostic category was associated with much higher NTA costs, $153.78, $138.79 and $101.67, respectively.

The variables that appeared in all three of the models have quite disparate effects across the acute, chronic, and rehabilitation NP categories, supporting the use of separate models for each category. As an example, tracheostomy care increased NTA costs by $15.11 to $35.56 per diem, depending on the NP category. The coefficients on the charge models (not shown) had similar patterns.

A few variables appear to be useful for assessing NTA costs that are not currently in the MDS, and these variables could be included in the MDS instrument. As noted in Chapter IV, one of these is the use of intravenous medication on both SNF claims and MDS data. It not only affects costs independently, but also appears to identify sicker patients within other services or conditions—appearing in a number of interaction variables, including IV use in conjunction with relief bed use—which increases the NTA cost by $55.74 in the chronic category. Accurately identifying the provision of respiratory therapy is also important. Simple changes in the MDS instrument could identify the delivery of IV services and the delivery of respiratory therapy in the SNF.

The model fit is shown in Table V.12 for both the cost and charge models, including both the R-squared and the percentage of the top 10% of cases (measured by charges and costs) in resource use accurately predicted in the top 10% by the model. The models explained 23%-25% of the variation in wage-adjusted costs and 35%-37% of wage-adjusted charges. The models were particularly adept at identifying high cost patients. Between 43.5% and 46.9% of stays in the top 10% of costs were predicted to be in the top 10% of costs, while all 3 charge models had more than one-half of stays with charges in the actual top 10% also in the predicted top 10%. These compare with 32% found in the Abt WIM models and 22% in RUG-III models (Abt Associates 2000).

Table V.12: Performance Summary for NP NTA Stay-Level Models

	Acute (N=27,504)	Chronic (N=24,310)	Rehabilitation (N=111,092)
Number of variables in model	21	22	29
Wage-adjusted costs			
R-squared	0.25	0.23	0.24
Percentage of top 10% of cost cases correctly predicted in top 10%	45.6	43.5	46.9
Wage-adjusted charges			
R-squared	0.37	0.35	0.36
Percentage of top 10% of charge cases correctly predicted in top 10%	55.1	51.2	53.5

Note: Models run on the test sample.

Validation and Facility-Level Analyses

We applied the coefficients derived from the models based on the test sample to the stays in the validation sample of facilities (described in Chapter III) as a way of testing the models. Performance summary statistics analogous to those in Table V.12 for the test sample models are shown for the validation sample models in the top portions of Table V.13 (for wage-adjusted cost models) and Table V.14 (for wage-adjusted charge models).

The validation procedure generated models with similar R-squared values and similar percentages of correct predictions of the cases in the top 10% of charges and costs per day compared to the original sample. For example, the wage-adjusted cost models for the validation sample have R-squared values of 0.23, 0.22 and 0.26, respectively for acute, chronic, and rehabilitation categories in Table V.13. The models for wage-adjusted charges show similar results in most cases, and the stay-level charge models relative to the cost models have, on average, higher R-squared values for the patient classification categories and do a slightly better job identifying high cost cases.

Table V.13: Performance Summary for Stay- and Facility-Level NTA Cost Models

	Acute	Chronic	Rehabilitation	Total sample
Stay-level Models				
Test Sample				
R-squared	0.25	0.23	0.24	0.25
Percentage of top 10% of cost cases correctly predicted in top 10%	45.6	43.5	46.9	46.1
STD of relative weights	0.70	0.67	0.67	0.68
Validation Sample R-squared	0.23	0.22	0.26	0.24
Percentage of top 10% of cost cases correctly predicted in top 10%	43.1	42.8	46.8	46.0
STD of relative weights	0.67	0.65	0.68	0.67
Facility Level R-squared	0.25	0.28	0.41	0.40
Facility-level Models (Validation Sample)				
CMI Only Model				
CMI Coefficient	1.05	1.08	1.14	1.15
R-squared	0.36	0.33	0.37	0.38
Standard Error	0.038	0.041	0.045	0.04
Coefficient different from 1	p=.18	p=.08	p<=.01	p<=.01
Payment Model				
CMI Coefficient	1.04	1.07	1.13	1.14
R-squared	0.37	0.33	0.37	0.38
Standard Error	0.038	0.04	0.043	0.04
Coefficient different from 1	p=.25	p=.10	p<=.01	p<=.01
Fully-Specified Model				
CMI Coefficient	0.93	0.94	0.98	0.99
R-squared	0.39	0.35	0.39	0.40
Standard Error	0.043	0.049	0.054	0.051
Coefficient different from 1	p=.11	p=.25	p=.70	p=.90

Notes: The test sample had the following sample sizes: Acute (N=27,504), Chronic (N=24,310), Rehabilitation (N=116,915), Total sample (N=162,906). The validation sample had the following sample sizes: Acute (N=28,262), Chronic (N=24,799), Rehabilitation (N=116, 915), Total sample (N= 169,976). Standard errors and P-values reported are the Huber-White standard errors, correcting for heteroskedasticity. The CMI models all run on sample of (N=1,384) facilities. The payment model includes urban/rural status and area wage index, in addition to the CMI. The fully-specified model includes urban/rural status, area wage index, SNF bed size indicators, hospital-based status, ownership type, and percent Medicare days, In addition to the CMI.

Table V.14: Performance Summary for Stay- and Facility-Level NTA Charge Models

	Acute	Chronic	Rehabilitation	Total sample
Stay-level models				
Test Sample				
R-squared	0.37	0.35	0.36	0.36
Percentage of top 10% of charge cases correctly predicted in top 10%	55.1	51.2	53.5	54.9
STD of relative weights	1.18	1.12	1.13	1.13
Validation Sample				
R-squared	0.33	0.31	0.34	0.34
Percentage of top 10% of charge cases correctly predicted in top 10%	51.2	49.1	52.9	49.1
STD of relative weights	1.14	1.10	1.10	1.14
Facility Level R-squared	0.41	0.39	0.59	0.50
Facility-level Models (Validation Sample)				
CMI-Only Model				
CMI Coefficient	0.50	0.55	0.55	0.57
R-squared	0.3	0.29	0.29	0.31
Standard Error	0.040	0.061	0.055	0.039
Coefficient different from 1	p<=.01	p<=.01	p<=.01	p<=.01
Payment Model				
CMI Coefficient	0.49	0.56	0.56	0.57
R-squared	0.32	0.34	0.33	0.34
Standard Error	0.036	0.058	0.051	0.036
Coefficient different from 1	p<=.01	p<=.01	p<=.01	p<=.01
Fully-Specified Model				
CMI Coefficient	0.42	0.46	0.46	0.47
R-squared	0.35	0.36	0.36	0.36
Standard Error	0.032	0.051	0.045	0.031
Coefficient different from 1	p<=.01	p<=.01	p<=.01	p<=.01

Notes: The test sample models had the following sample sizes: Acute (N=27,504), Chronic (N=24,310), Rehabilitation (N=116,915), Total sample (N=162,906). The validation sample had the following sample sizes: Acute (N=28,262), Chronic (N=24,799), Rehabilitation (N=116, 915), Total sample (N= 169,976). Standard errors and P-values reported are the Huber-White standard errors, correcting for heteroskedasticity. The CMI models all run on sample of (N=1,384) facilities. The payment model includes urban/rural status and area wage index, in addition to the CMI. The fully-specified model includes urban/rural status, area wage index, SNF bed size indicators, hospital-based status, ownership type, and percent Medicare days, in addition to the CMI.

As a further validation step, we "rolled up" the predictions to derive an average predicted per diem NTA resource use over all Medicare covered days at each facility. The facility-level R-squared values from regressions on the association between the average predicted per diem NTA resource use and the actual per diem NTA resource use over all Medicare covered days at each facility provide a measure of model fit at the facility level and are shown in the seventh row of Tables V.13-V.14. For example, the cost models (Table V.13) yield R-squared values of 0.25, 0.28, and 0.41, respectively, for the acute, chronic, and rehabilitation NP categories, and the charge models (Table V.14) yield somewhat higher values of 0.41, 0.39, and 0.59.

Finally, we joined stays in the three NP patient classification categories to examine the model performance on the entire sample. The results are in the last column of Tables V.13-V.14. For example, taking the predictions from each of the three cost models and comparing them with the actual per diem costs in Table V.13 results in an overall fit of 0.25, and correctly predicts 46.1% of the cases with the top 10% of costs. (The corresponding figures for the charge models in Table V.14 are 0.36 and 54.9%.)

Facility Case Mix Index (CMI) Models

As discussed in Chapter III, a CMI is a way of measuring facility differences in patient mix that lead to different average NTA cost burdens (and therefore expected payments). To examine potential CMI coefficient compression, we estimated the three specifications of CMI models outlined in Chapter III: the CMI-only model, the payment model and the fully-specified model. (Please refer to Chapter III for a full discussion of the CMI methodology.) Both cost-based and charge-based CMIs were developed. The results are shown in the "facility-level" sections of Tables V.13 and V.14 (for cost models and charge models, respectively).

The cost-based CMI models (Table V.13) show some evidence of compression, with coefficients on the CMI-only model between 1.05 and 1.15. All of the models that included all facility characteristics, however, have coefficients that are not significantly different from one. The fit for all of the models is quite good, with the total sample model having an R-squared value of .38 in the CMI-only model and 0.40 in the fully-specified model. However, in the models based on charge-based CMIs, the facility-level cost regressions show decompression; that is, they generated case mix differences among facilities that were larger than the actual variation in resources used. For example, we find a CMI coefficient of 0.47 for the total sample in the fully-specified model using charge-based CMIs (Table V.14) compared to a coefficient of 0.99 using cost-based CMIs (Table V.13).

Charge-based CMIs also exhibit a broader distribution than cost-based CMIs. This is consistent with the higher level of variation in the charge-based relative weights compared with the cost-based relative weights, possibly because charge CMIs may retain more variation among facilities than the average facility costs exhibit. Also, although not presented here, coefficients from fully-specified models provide some insight into cost and charge patterns by facility characteristics. For example, rural facilities had significantly higher NTA costs than urban facilities. Such information could be useful in considering payment adjustments that might be appropriate for certain facility types (e.g., rural). Overall, NPA facility-level analyses suggest that basing NTA payments on the NP patient classification system developed in this chapter would pay facilities roughly in proportion to NTA costs.

D. Discussion

The NP-NTA model presented here is based on clinically relevant patient characteristics and specially identified service indicators, rather than on more gameable characteristics or services. The models performed well in explaining the observed variation in costs and charges at both the stay and facility levels. Further, the models identified nearly half of the high cost stays in a sensitivity test. The CMI analyses showed a small level of compression for the cost models, but this disappeared in the presence of facility characteristics thought to influence facility costs.

This classification model system represents a significant change from the current system of no separate payment system for NTA costs. CMS data and payment calculation systems would need to be modified to incorporate the new payment component. Some additional data would be needed from hospital information and from SNF claims or possibly new MDS items. The NP-NTA models rely on information from the prior hospital stay. While there are barriers to transferring information from hospital to SNF in a timely manner, the advances in information technology and the limited amount of information required suggest that such barriers can be overcome. In addition, many of the hospital variables, such as discharge diagnoses, would be valuable information for SNF caregivers in and of themselves. And the deadline for the information transfer could be as late as the deadline for the first MDS submission. In addition, a few key variables in the NP-NTA model (such as IV use and respiratory therapy) currently must be obtained from SNF claims data. New MDS variables that accurately identify the required information would be preferable but would require time to design, test, and implement. Thus, both time and effort would be required by both CMS and the SNFs for development, implementation, and ongoing operational steps. In addition, since SNFs have become comfortable with the current system, they may be wary of changing to a new one—even if it promises to more accurately compensate them for the costliness of their patients.

E. NP-Therapy Models

As discussed in our prior report (Maxwell, et al. 2003) and mentioned above, a unique aspect of the SNF PPS is that it essentially uses a fee schedule-like mechanism to pay for rehabilitation therapy, through basing rehabilitation therapy payment on the number of ordered therapy minutes. A key goal of our analyses in this project is to explore whether SNF therapy costs can be predicted using mainly patients' clinical and functional characteristics. Similar exploratory efforts were not successful in the original RUG-III development work (Fries, et al. 1994). But this lack of success was due, at least in part, to the small number of Medicare SNF stays available in that study sample, and the heterogeneous combination of Medicare SNF patients and Medicaid nursing home residents in that sample. In this study, we have the opportunity to explore classification options that cater to the Medicare SNF population only, rather than all nursing home patients.

A cautionary note is needed, however, before proceeding. Evidence exists that facilities' behavior with respect to therapy use changed in response to PPS, and that these changes may make it more difficult to find relationships between patient characteristics and therapy cost. For example, Hutt, et al. (2001) found that facilities participating in the PPS demonstration increased the provision of therapy relative to its provision prior to the demonstration, despite finding no changes in resident case mix. Furthermore, this increase in therapy resulted in no improvement in community discharge rates. As PPS was being phased in, Kramer, et al. (2002) found steadily decreasing percentages of residents in the ultra high and very high rehabilitation RUGs—the high use categories—and steadily increasing percentages of residents in the high and moderate rehabilitation RUGs. They also found clustering of rehabilitation patients in a few RUG categories that received add-on payments. Because patient characteristics can more easily predict large differences in therapy payment than minor differences, this clustering of patients into selected RUG categories associated with small differences in therapy cost makes it more difficult to identify patient predictors of therapy cost. This is supported by White (2003), who noted that the proportion of SNF residents receiving very high or very low amounts of therapy dropped from 1997 to 2000, while the percentage of residents receiving moderate amounts of therapy increased over the same pre- and post-PPS period. We found that in the for-profit SNFs, where this was most likely to occur, patient characteristics were less predictive of therapy costs than in other types of SNFs.

Taken together, these findings suggest that predicting current SNF charges (or costs) may result in a model that reflects the particular incentives at work in the current reimbursement system, not necessarily patient need, and that our modeling may be hampered by this disconnect. Nonetheless, current charges and costs are the best data available on a current and repeated basis. With payments based initially on average costs, it is to be hoped that better incentives in the payment system will lead average costs to converge over time more closely to the efficient marginal cost of providing high quality care.

In this section, we first describe the methods pertaining specifically to payment model development for therapy services. The overall approach is similar to that used for NTA services and covered in the preceding sections of this chapter. Results are then presented relating to (a) distributions of costs and charges for therapy services, (b) our preliminary models using charges as the dependent variable, and (c) our more refined cost models at both the stay and facility levels. These results are followed by a discussion of the issues relating to payment system development for therapy services.

Methods

Following the NTA analysis approach, the test sample and validation sample described in Chapter III were used for these analyses, with charges and (wage-adjusted) costs per day as the dependent variables. Our objective was to identify factors that were associated with changes in therapy resource use (charges and costs per day) and could be incorporated into a payment system. We developed models for each NP group (acute, chronic, and rehabilitation) separately and for all three groups combined.

Analyses were run separately for the different components that constitute total therapy charges and costs. Physical therapy (PT) and occupational therapy (OT) had similar utilization patterns, had similar relationships with other variables, and were provided to similar patients; these components were summed, therefore, to create a combined PT/OT per diem charge and a combined PT/OT per diem cost. Speech therapy (ST) was in sharp contrast—with a quite different pattern of utilization, dissimilar relationships with other variables, and distinctly different types of patients to whom it was provided. As such, ST was analyzed separately.

The analyses began with an examination of distributions of charges and costs (only the findings for charges are presented since those for costs were similar). Then, correlations between dependent and possible independent variables were examined and stepwise regressions were run to identify and refine the group of independent variables to be included in the final models. Facility characteristics and certain resident characteristics that would be inappropriate for use in a payment system (e.g., bedfast, use of feeding tube) were not entered into the models regardless of the strength of their association with charges. Key variables were identified in terms of their estimated impact on charges per day. For dichotomous variables, the dollar amount is simply the variable's coefficient. For continuous variables, we calculated the effect on charges by multiplying the coefficient by the average value of the continuous variable.

We conducted further analyses using wage-adjusted per diem costs as the dependent variable instead of per diem charges. We examined correlations between possible independent variables and the cost variables, and also employed regression tree modeling (using CART) to identify potentially important interactive and nonlinear variables.[21] Refined stay-level models were estimated and performance statistics generated as was done in the NTA analyses. Thereafter, validation analyses were carried out and facility-level models were estimated using the validation sample. (The facility-level models were the CMI-only, payment, and fully-specified models described in Chapter III and the previous sections of this chapter.) Model statistics including the R-squared, coefficient of logged CMI, standard error, and p-value of a test for CMI coefficient = 1.0 were generated.

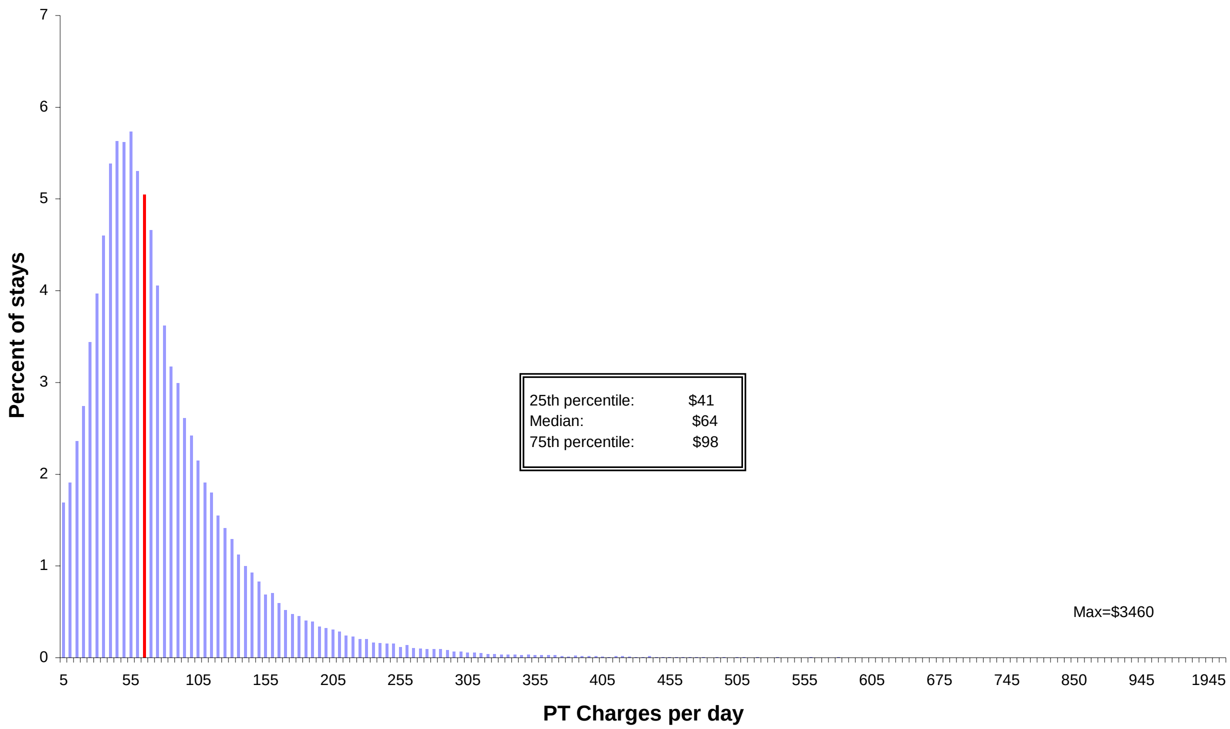
[21] Bedfast was particularly problematic in that it was strongly associated with therapy costs but including it in a payment system would generate an undesirable incentive. We therefore modeled bedfast as a dependent variable and created proxy bedfast variables consisting of combinations of mobility ADLs (bed mobility, transfer, walk in room, walk in corridor, and locomotion on and off unit). These proxy bedfast variables were included in the stepwise regressions if they were correlated with PT/OT or ST costs at $R \geq 0.10$.

F. Results

Distributions

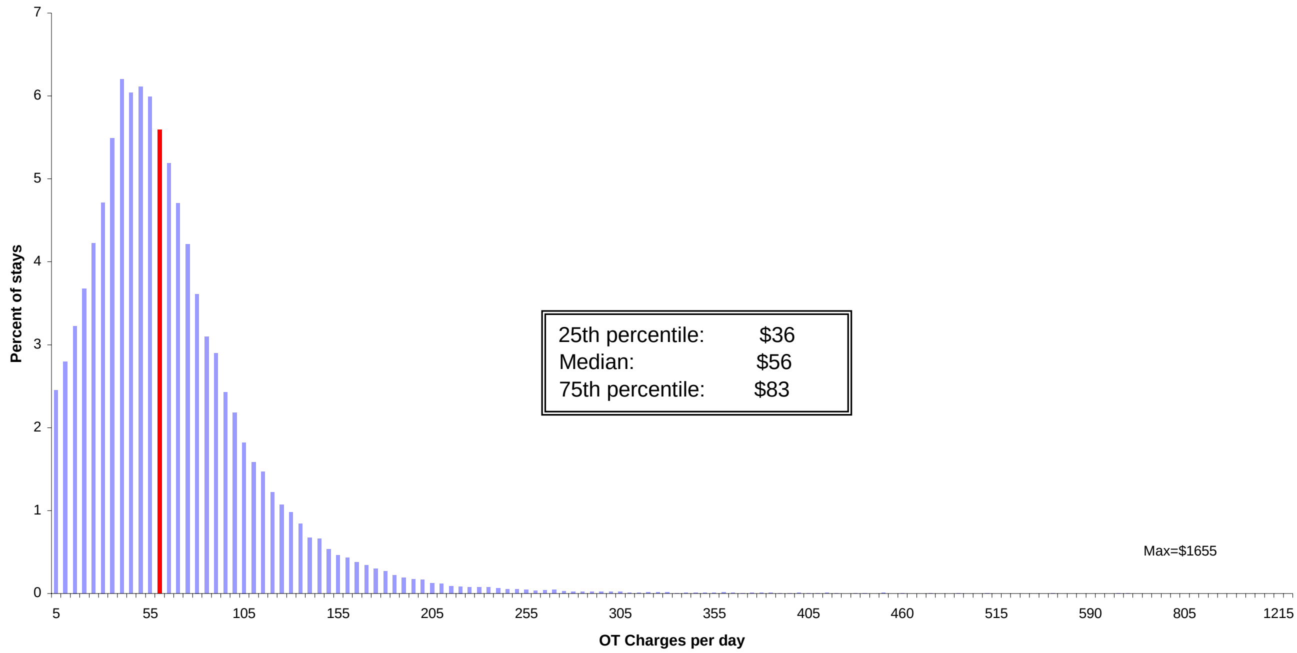
Distributions of charges for PT and OT are provided in Figures V.7 and V.8, respectively, for patients who received at least some of the service. Both of these distributions are bell shaped, but with a tail to the right representing the stays with substantial PT and OT use. The left side of both distributions drops sharply, because only a small proportion of patients received minimal PT or OT services. The median value for PT is $8 per day higher than for OT, but the use pattern of these 2 services is very similar. The maximum value for both these services was extremely high prior to eliminating outliers.

Figure V.7: PT Charge Distribution (All Stays with PT Charges > $0)



Note: Distribution is based upon the test sample, prior to application of data exclusion rules, and illustrates charges per day for 140,645 stays with more than $0 PT charges

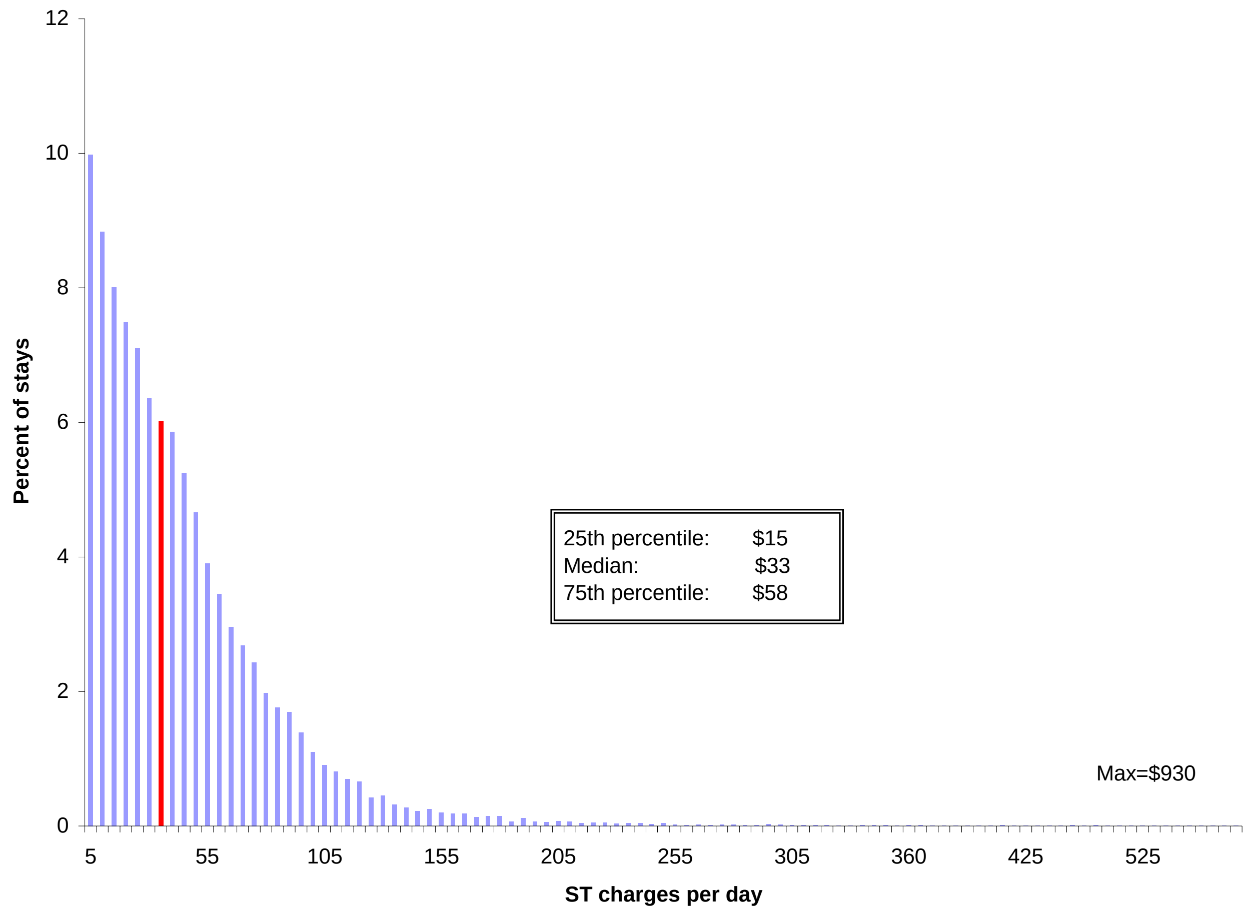
Figure V.8: OT Charge Distribution (All Stays with OT Charges > $0)



Note: Distribution is based upon the test sample, prior to application of data exclusion rules, and illustrates charges per day for 123,284 stays with more than $0 OT charges

The distribution for ST (Figure V.9) is entirely different from the other two distributions. Of the people receiving ST services, the vast majority received relatively small amounts per day, with a median daily charge of $33. The use pattern is quite different from those of PT and OT, with a declining number of users as charges per day increase. High-end users still exist, but are not as extreme as for PT and OT. This distinctly different use pattern suggests that ST use may be driven by a different process, making it worthwhile to examine ST separately from PT and OT.

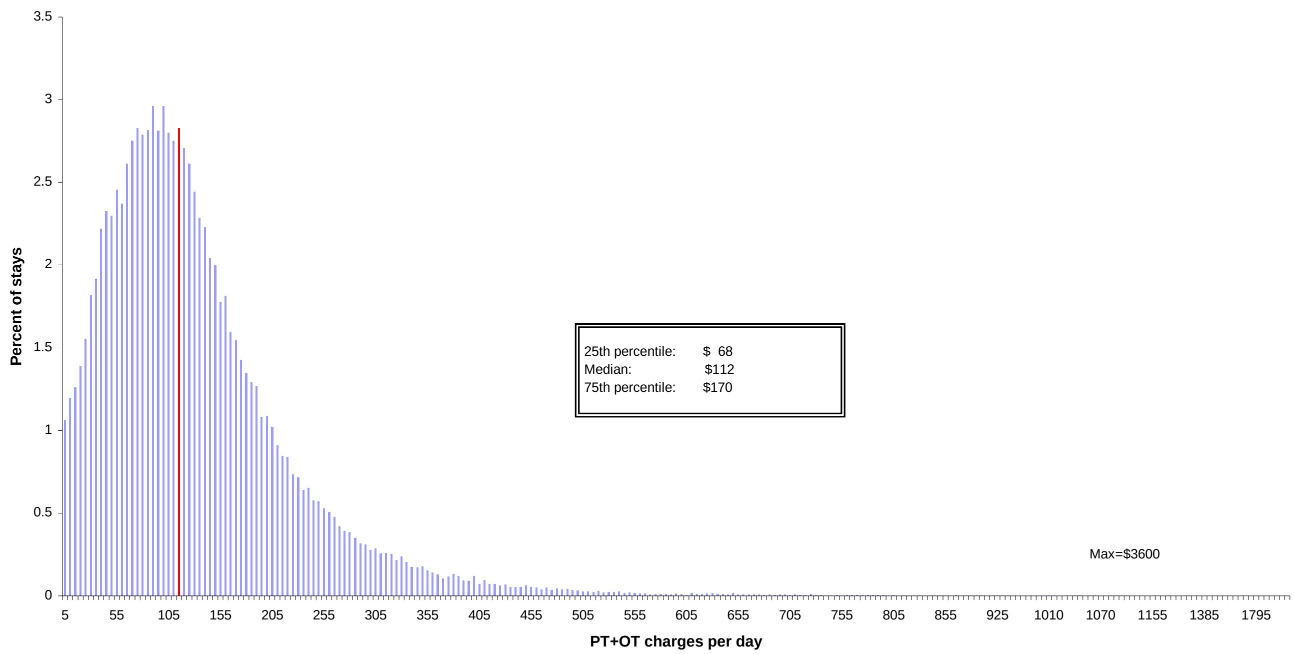
Figure V.9: ST Charge Distribution (All Stays with ST Charges > $0)



Note: Distribution is based upon the test sample, prior to application of data exclusion rules, and illustrates charges per day for 32,879 stays with more than $0 ST charges

Figure V.10 shows the distribution of the sum of PT and OT charges. As expected, the curve looks similar to the individual PT and OT curves.

Figure V.10: PT/OT Charge Distribution (All Stays with PT/OT Charges > $0)



Note: Distribution is based upon the test sample, prior to application of data exclusion rules, and illustrates charges per day for 144,729 stays with more than $0 PT/OT charges

Also important to note is that the therapy distributions show a high percentage of cases with zero therapy charges and some very high per-day charges in the right-hand tail of each distribution. Table V.18 shows that over 80% of SNF patients have PT charges, nearly three-quarters have OT charges, and about 85% have PT or OT charges. In contrast, less than 20% receive ST, which is a distinguishing characteristic of ST relative to PT and OT. The 95th percentile is about 3 times the median for the different therapy components. (The medians are shown in Figures V.7-V.10.)

Table V.18: Distribution Skew for Therapy Charges

	% w/ no charges	95th percentile	97th percentile	99th percentile
PT	17.7	$189	$222	$301
OT	27.8	146	168	221
ST	80.8	114	133	189
PT/OT	15.3	306	352	466

Note: Statistics are based upon the test sample, prior to application of data exclusion rules. Percentiles are based on only those stays with more than $0 charges by discipline (n=140,645 for PT, n=123,284 for OT, n=32,879 for ST, n=144,729 for PT/OT).

Variation in therapy costs and charges per day is depicted in Table V.19 for all stays and the three NP groups. As expected, mean PT/OT charges are higher for the rehabilitation group than for the acute and chronic groups. The standard deviation is similar for all three groups; thus, the coefficient of variation (CV), representing the ratio of the standard deviation to the mean, is lowest for the rehabilitation group relative to acute and chronic groups. The CV provides a relative comparison of variation in which lower CVs represent less variation. The reduced variation in the rehabilitation group probably represents a narrowing of the distribution of therapy use under the current PPS, in which there is an incentive to provide specific levels of therapy for those receiving rehabilitation. As expected, average PT/OT costs are substantially lower than charges following application of the cost-to-charge ratios. However, the means, standard deviations, and CVs follow a similar pattern for costs as they did for charges across the three patient categories.

Table V.19: Variation in Therapy Cost and Charges Per Day

	All Stays (n=163,738)	Acute (n=27,563)	Chronic (n=24,392)	Rehabilitation (n=111,184)
PT/OT Charges				
Mean	113.27	89.74	80.20	126.49
Standard Deviation	99.86	92.95	96.62	99.53
Coefficient of Variation	0.88	1.04	1.20	0.79
PT/OT Costs				
Mean	64.88	52.08	45.97	75.24
Standard Deviation	51.33	50.03	50.20	50.19
Coefficient of Variation	0.79	0.96	1.09	0.69
ST Charges				
Mean	8.43	12.27	8.59	7.45
Standard Deviation	25.14	30.73	25.19	23.45
Coefficient of Variation	2.98	2.51	2.93	3.15
ST Costs				
Mean	5.20	7.67	5.51	4.51
Standard Deviation	15.62	19.48	16.21	14.21
Coefficient of Variation	3.01	2.54	3.00	3.15

Note: Statistics are based upon the test sample

Mean speech therapy charges and costs are substantially less than the respective means for PT/OT, as expected with the huge number of zero values for ST. However, as indicated earlier, when ST is provided, the cost is significant. Due to the high standard deviation resulting from large numbers of zeros and fairly high values on the other extreme, the standard deviation and coefficient of variation are both substantial for ST charges and costs.

Preliminary Models

In selecting predictors for therapy charges, we began by examining correlations separately for PT/OT and ST charges. We reasoned that if correlations were similar for PT and

OT relative to ST, then we could combine all 3 therapy services, despite the differences in distributions. However, as indicated in Table V.20, we found opposite correlations between PT/OT charges and ST charges for many of our key variables. For example, low-Barthel Index representing extreme dependence was associated with less PT/OT but higher ST charges. The mid-Barthel Index, associated with the highest PT and OT charges, was actually negatively associated with ST charges. The Cognitive Performance Scale (CPS), in which higher values were associated with greater cognitive impairment, was negatively associated with PT/OT charges, suggesting less therapy for those who are cognitively impaired, but positively associated with the ST charges. Individual ADLs follow the same pattern as the Barthel Index, with greater independence on the eating ADL scale associated with less PT and OT but more speech therapy. As expected, swallowing problems, which are treated by speech therapists, were associated with higher ST charges but with lower PT and OT charges. Even conditions and diagnoses like surgical wound care and musculoskeletal diagnosis have opposite correlations between PT/OT and ST.

Table V.20: Correlations of Predictors with PT/OT vs. ST Charges

	PT/OT Charges	ST Charges
Low Barthel Index	-0.240	+0.122
Mid Barthel Index	+0.193	-0.053
Eating Assistance Required	-0.263	+0.187
Eating (independent)	+0.188	-0.168
Cognitive Performance Impaired	-0.253	+0.075
Cognitive Performance Intact	+0.223	-0.160
Swallowing Problem	-0.133	+0.319
Surgical Wound	+0.186	-0.073
Musculoskeletal Diagnosis	+0.140	-0.065

Note: Statistics are based upon the test sample, prior to application of data exclusion rules

Such a pronounced difference in correlation structure between predictors and the dependent variable indicates the usefulness of examining the two components separately. Thus, our strategy was to build models separately for PT/OT charges and ST charges, with the idea that these could be combined in building the therapy payment model. In the next stage of analysis, we plan to explore the models that combine the two components.

Models were built using variables from the qualifying hospital stay and the MDS at the resident level. These models did not include any facility characteristics, which would be considered after aggregating to the facility level. Person-level R-squared for the initial resident-level charge models are provided in Table V.21 separately for PT/OT and ST. Our ability to explain PT/OT variance was less in the rehabilitation group than for either the acute or chronic group. We were also less successful in explaining variation in ST use among patients in the chronic group. Interestingly, among the three NP groups, the chronic group had models that explained the highest amount of PT/OT variations and the lowest amount of ST variation.

Table V.21: Explanatory Power for Initial Charge Models

	PT/OT	ST
Acute	0.248	0.190
Chronic	0.337	0.113
Rehabilitation	0.164	0.209

Note: Statistics are based upon the test sample, prior to application of data exclusion rules

Variables associated with at least $5 a day in PT/OT charges based on the regression results are provided in Table V.22, with the dollar value of this association indicated for each patient category for which the association is significant. Qualifying hospital variables are listed first, followed by MDS variables. These initial models included only main effects without interaction terms. Charges per day in the hospital relating to both therapies and total charges were highly associated with increased PT and OT use particularly for the rehabilitation group. Several diagnoses related to medical problems were negatively associated with PT/OT use and those related to musculoskeletal disease were positively associated with PT/OT as expected. For a number of the ADLs (e.g., bed mobility), the moderate level of function was associated with higher PT/OT therapy use than the most dependent level of function using independent function as the reference.[22] In some cases (e.g., walking in room and in corridor), however, a more linear relationship appears to exist between ADLs and PT/OT use, with the most dependent levels

[22] For the PT/OT analyses, the ADL scales were transformed into dichotomies representing moderate (mod) and high (max) dependency levels. The purpose of this approach was to reflect the non-linear nature of the relationship between ADLs and PT/OT use.

associated with the highest levels of PT/OT use. Neither the variables nor the coefficients for the variables were consistent across the three NP groups, reinforcing the need to continue with separate models initially.

Table V.22: Variables Associated with at least $5 of PT/OT Charges per Day

Variable	Acute (n=28,600)	Chronic (n=25,700)	Rehab (n=114,908)
Qualifying Hospital Variables			
Hospital - Respiratory infection	-10.18		
Hospital - Mental disorders	-7.28		
Hospital –Skin/subcutaneous tissue disease		-9.88	
Hospital – Osteoarthritis		12.05	10.19
Hospital - Mental disorders			-6.03
MCD - Infectious & parasitic	-12.75	-7.06	
MDC - Musculoskeletal system			10.11
PT/OT charges per day prior hosp'n	14.11	14.00	17.87
Total charges per day prior hosp'n	11.62	8.47	16.71
Potential acute diagnosis		-6.82	
MDS Variables			
ADLs			
Bed mobility (mod)	10.81	15.21	
Bed mobility (max)	8.33	13.54	
Transfer (mod)	16.41	9.39	
Transfer (max)	5.78		
Walk in room (mod)	9.15	9.48	
Walk in room (max)	28.47	16.98	
Walk in corridor (mod)	7.10	7.58	
Walk in corridor (max)	14.99	15.39	
Locomotion on unit (mod)	21.00	20.92	
Locomotion on unit (max)	12.92	14.05	
Locomotion off unit (mod)	6.08	9.15	
Locomotion off unit (max)	23.24	24.79	
Dressing (mod)	8.01	12.65	
Dressing (max)	7.87	12.92	
Eating (mod)		-6.88	8.10
Eating (max)	-15.35	-21.90	
Toileting (mod)	11.47	10.46	
Toileting (max)	13.81	5.56	
Hygiene (mod)	-6.06	-11.14	
Hygiene (max)	-21.62	-23.34	-8.62
Bathing (mod)		-6.64	14.40
Bathing (max)	-23.61	-19.43	
Number of 6 ADLs that are max dep	-31.66	-28.46	
5 day Barthel Index			22.96
Others			
Cognitive skills scale	-14.40	-12.47	-8.99
Surgical wound		10.38	14.07
Charges R-squared	0.248	0.337	0.164

Note: Statistics are based upon the test sample, prior to application of data exclusion rules.

Variables associated with at least $1 per day in ST charges are provided in Table V.23. Qualifying hospital variables include ST charges per day and diagnoses related to stroke and nervous system, which were expected to be associated with increased ST charges. The MDS ADL scales were generally associated with lower ST charges. Swallowing problems, memory problems, and chewing problems were all associated with higher ST charges. Once again, differences were found among the three NP groups.

Table V.23: Variables Associated with at least $1 of ST Charges per Day

Variable	Acute (n=28,604)	Chronic (n=25,703)	Rehab (n=114,908)
Qualifying Hospital Variables			
Hospital - Stroke	14.33	8.27	11.31
MDC - Nervous system	1.45	1.97	1.51
ST charges per day prior hosp'n	2.36	1.10	1.29
MDS Variables			
ADLs			
Bed mobility scale	-2.58	-3.05	
Transfer scale	-2.49	-2.42	
Walk in room scale		2.50	
Walk in corridor scale		-2.69	
Locomotion on unit scale		1.68	
Toileting scale		-1.70	
Hygiene scale	-3.19	-1.47	
Bathing scale		-1.73	
All 6 mobility ADLs are max dep	2.19		
5 day Barthel Index	-5.46	-6.89	
Others			
Memory problem	3.07	2.36	2.18
Cognitive skills scale	-2.18	-1.69	
Turn/reposition program	-1.08		
Swallowing problem	12.28		17.61
Chewing problem	1.62	1.14	1.67
Charges R-squared	0.190	0.113	0.209

Note: Statistics are based upon the test sample, prior to application of data exclusion rules.

Final Therapy Cost Models

More refined models were generated separately for PT/OT costs and ST costs using wage-adjusted costs as the dependent variable rather than charges. These stay-level models were then validated on the validation sample in which, as noted, facilities were randomly selected and all stays within each facility were included. Facility-level R-squared and CMIs were then calculated using these final models.

Table V.24 provides the means for all variables in the final PT/OT and ST cost models. These are separated into the qualifying hospital variables, MDS variables (distinguishing ADLs from others), and interaction terms. Qualifying hospital variables include the PT/OT and ST charge per day from the previous hospital stay and a series of diagnoses or specific major diagnostic categories (MDCs).

Table V.24: Means for Variables in the Final PT/OT and ST Cost Models

Variable	Mean (% unless indicated)
Qualifying Hospital Variables	
Hospital - Hip fracture	9.0 %
Hospital - Dementia	20.4
Hospital - Injury	28.1
Hospital - Mental disorders	33.2
Hospital - Musculoskeletal & connective tissue disease	31.9
Hospital - Osteoarthritis	12.9
Hospital - Respiratory infection	17.3
Hospital - Stroke	7.6
Hospital - Skin/Subcutaneous tissue disease	13.1
MDC - Infectious & parasitic	16.7
MDC - Musculoskeletal system	44.5
MDC - Respiratory system	41.9
MDC - Skin, subcutaneous tissue	15.7
MDC - Nervous system	37.2
Potential acute diagnosis	72.7
PT/OT charges per day ($)	81.50
ST charges per day ($)	8.92
Speech therapy use	15.1
MDS Variables	
ADLs	
Bed mobility (mod)	50.7
Bed mobility (max)	25.6
Transfer (mod)	57.7
Transfer (max)	32.0
Walk in room (mod)	65.8
Walk in room (max)	24.2

Table V.24: Means for Variables in the Final PT/OT and ST Cost Models

Variable	Mean (% unless indicated)
Walk in corridor (mod)	65.6
Walk in corridor (max)	26.7
Locomotion on unit (mod)	43.9
Locomotion on unit (max)	41.6
Locomotion off unit (mod)	35.3
Locomotion off unit (max)	54.8
Dressing (mod)	56.1
Dressing (max)	36.5
Eating (mod)	22.9
Eating (max)	17.0
Toileting (mod)	51.5
Toileting (max)	38.8
Hygiene (mod)	52.3
Hygiene (max)	32.6
Bathing (mod)	76.5
Bathing (max)	11.8
Low 5-day Barthel Index	19.1
Others	
Memory problem	49.5
Cognitive skills (mid)	24.9
Cognitive skills (mod-sev impair)	21.0
Making self understood (usually/sometimes)	22.5
Making self understood (rarely/never)	4.5
Speech clarity (unclear/none)	13.6
Understand others (usually)	18.2
Understand others (sometimes/rarely/never)	14.6
Chewing problem	16.7
Swallowing problem	17.4
Surgical wound	28.4
IV med/soln from SNF claim or MDS	9.1
Interaction Variables	
Stroke & cognitive skills (any impair)	5.7
Stroke & eating (any dep)	5.5
Stroke & swallowing problem	3.4
Swallowing problem & cognitive skills (any impair)	15.4
Swallowing problem & cognitive skills (mod-sev impair)	4.9
Swallowing problem & eating (any dep)	16.3
Surgical wound & arthritis	14.7
Eating (min/mod) & surgical wound	12.5

Note: Figures are based upon the test sample and are percentages unless otherwise noted

Table V.25 provides the regression coefficients for all variables present in the final PT/OT cost models, by patient group. Hospital charges are positively associated with PT/OT costs in the SNF. While diagnoses such as osteoarthritis, musculoskeletal system, injury, and hip fracture are positively associated with PT/OT in various groups, the medical diagnoses are typically negatively associated with PT/OT, as we might expect. As previously discussed, the MDS variables follow two patterns. Many (e.g., bed mobility, transfer, locomotion on unit) are associated with increased PT/OT use at the moderate dependency level, with a decrease or a smaller increase in PT/OT cost at the more dependent level. Other ADLs (e.g., walking in room, walking in corridor) have a more linear relationship with costs, in which the moderate level is associated with a moderate increase in costs and the dependent level is associated with a higher increase in cost. Once again, these associations reinforce the need to treat ADLs in a non-linear fashion when predicting PT/OT costs. A number of other MDS variables found to be predictive of PT/OT costs have also been added to the file. These include cognitive skills, understanding others, and making oneself understood, as well as problems related to swallowing and chewing. IV medication was associated with less PT/OT, not surprisingly since such individuals are unlikely to be able to participate fully. Surgical wounds, particularly in the interaction terms of eating and arthritis, were predictive of PT/OT costs.

Table V.25: Final PT/OT Models: Variables and their Dollar Effects on Cost per Day

Variables	Acute (n=27,104)	Chronic (n=24,083)	Rehabilitation (n=110,977)
Qualifying Hospital Variables			
Hospital - Respiratory infection	-7.37		
Hospital - Dementia	-3.99		-3.42
Hospital - Injury	4.92		
Hospital - Skin/Subcutaneous tissue disease	-9.46	-8.38	
Hospital - Osteoarthritis		7.06	9.05
Hospital - Mental disorders		-2.43	
Hospital - Musculoskeletal & connective tissue disease			-1.61
Hospital - Hip fracture			7.82
MDC - Infectious & parasitic	-7.61	-4.20	
MDC - Skin, subcutaneous tissue		2.13	
MDC - Musculoskeletal system		2.67	4.57
MDC - Respiratory system		-1.94	
PT/OT charges per day	0.10	0.10	0.08
Potential acute diagnosis		-3.37	
MDS Variables			
ADLs			
Bed mobility (mod)	2.58	2.70	2.97
Bed mobility (max)	-0.93	-1.69	5.33
Transfer (mod)	8.56	6.00	8.51
Transfer (max	2.83	0.76	4.78
Walk in room (mod)	3.24	2.75	6.52
Walk in room (max)	13.50	8.49	7.19
Walk in corridor (mod)	6.85	5.08	2.47
Walk in corridor (max)	12.14	6.05	1.10
Locomotion on unit (mod)	6.42	7.48	
Locomotion on unit (max)	2.84	5.36	
Locomotion off unit (mod)	-0.27	2.80	
Locomotion off unit (max)	3.34	5.09	
Dressing (mod)	3.99	6.99	2.14
Dressing (max)	3.90	6.81	6.35
Eating (mod)	-0.46	-5.53	-1.57
Eating (max)	-7.98	-13.62	-9.61
Toileting (mod)	4.20	2.20	3.63
Toileting (max)	5.74	-3.04	1.37
Hygiene (mod)	-1.88	-4.95	-2.77
Hygiene (max)	-8.12	-10.77	-11.04
Bathing (mod)	1.80		0.25
Bathing (max)	-7.13		-6.54
Low 5-day Barthel Index	-13.71		
Others			
Cognitive skills (mid)	-3.22	-3.16	-2.24
Cognitive skills (mod-sev impair)	-8.70	-8.22	-3.17
Understand others (usually)	-1.79		-1.21
Understand others (sometimes/rarely/never)	-7.46		-4.99
Making self understood (usually/sometimes)		-0.14	
Making self understood (rarely/never)		-8.88	
Memory problem			-4.06
Swallowing problem	2.71	1.95	
Chewing problem		-1.62	
Surgical wound		2.12	-1.07
IV med/soln from SNF claim or MDS	-11.45		
Interaction Variables			
Eating (min/mod) and surgical wound		14.55	
Surgical wound and arthritis			19.12
Wage-Adjusted Costs R-squared	0.191	0.269	0.121

Note: Statistics are based upon the test sample.

In the final PT/OT cost models, the person-level R-squared values were lower for the rehabilitation group than the acute and chronic groups, probably because of the effects of PPS previously mentioned. Therapy amounts and hence therapy cost for patients in rehabilitation RUGs appear to be driven by payment policy, resulting in less cost variation and variation that is less tied to patient characteristics. To the extent that the lack of variation in therapy costs is an artifact of the existing payment system, in which therapy is provided according to payment category incentives rather than specific patient needs, it will be difficult to relate patient characteristics to therapy costs.

Table V.26 includes the variables and their regression coefficients for the final ST cost models. Two prior hospital use variables for ST were highly predictive of ST costs in the SNF, including the dichotomous indicator of whether any speech therapy was provided and the charge per day variable. Two diagnoses that we would expect to be associated with ST costs were also in the model, stroke, and nervous system diseases. Although the ADLs were generally less strongly associated with ST, they typically followed the pattern of increased costs at the moderate dependence level and either smaller increases or decreased costs at the most dependent level, as we found earlier. The variables relating to swallowing, chewing, cognition, and speech and language were all associated with ST, as expected. Scales in some of these instances were broken into dichotomies (e.g., making self understood) so that the mid-range was associated with increased ST costs and the most dependent range was associated with decreased costs. Several interaction terms were also important, particularly in the rehabilitation group. The R-squared for these models did not change much from the charge models.

Table V.26: Final ST Models: Variables and Their Dollar Effects on Cost per Day

Variables	Acute (n = 27,098)	Chronic (n = 24,059)	Rehabilitation (n = 110,955)
Qualifying Hospital Variables			
Hospital - Stroke	8.11	5.39	2.82
MDC- Nervous system	0.99	1.58	1.00
Speech therapy use	4.98	3.81	4.16
Speech therapy charge per day	0.06	0.05	0.05
MDS Variables			
ADLS			
Bed mobility (mod)		0.62	
Bed mobility (max)		-1.08	
Walk in corridor (mod)		-0.68	
Walk in corridor (max)		1.13	
Eating (mod)	1.99	2.30	0.60
Eating (max)	0.62	1.22	0.32
Hygiene (mod)			0.24
Hygiene (max)			-0.13
Bathing (mod)		0.10	
Bathing (max)		-1.26	
Others			
Swallowing problem	7.30	8.55	7.13
Chewing problem	1.46		0.92
Cognitive skills (mid)	1.46	0.94	0.67
Cognitive skills (mod-sev impair)	0.75	1.07	0.43
Making self understood (usually/sometimes)	1.99	1.58	1.47
Making self understood (rarely/never)	-3.04	-2.41	-0.83
Speech clarity (unclear/none)	2.65	1.14	2.64
Understand others (usually)	-0.17	0.72	-0.10
Understand others (sometimes/rarely/never)	-2.57	-0.05	-0.92
Memory problem			1.13
Interaction Variables			
Stroke & swallowing problem			3.93
Stroke & eating (any dep)			2.04
Stroke & cognitive skills (any impair)			1.48
Swallowing problem & cognitive skills (any impair)			0.97
Swallowing problem & cognitive skills (mod-sev impair)		-4.20	
Swallowing problem & eating (any dep)			2.81
Wage-Adjusted Costs R-squared	0.185	0.120	0.207

Note: Statistics are based upon the test sample.

Validation and Facility-Level Analyses

Table V.27 provides both stay and facility-level model fit using wage-adjusted costs as the dependent variable for PT/OT. In the first line of the table, the R-squared values for the patient stay models for each NP group from Table V.25 are provided, along with the R-squared of the total sample (0.192). The model is stable, as reflected by the validation R-squared values, which are comparable. The proportion of stays with costs in the top 10% that are predicted by the model to be in the top 10% is also relatively comparable in the two samples (i.e., 27.7% in the test sample, 29.3% in the validation sample).

The R-squared at the facility level drops considerably for the rehabilitation group and the total sample. Although the variation in PT/OT cost did decrease after aggregating to the facility level (the standard deviation developed from 52 at the resident level to 32 at the facility level), the amount of smoothing was not nearly as large as that seen in NTA costs (where the standard deviation declined from 82 to 35).

Table V.27: Performance Summary for Stay and Facility-Level PT/OT Cost Models

	Acute	Chronic	Rehabilitation	Total sample
Stay-level models				
Test Sample				
R-squared	0.191	0.269	0.121	0.192
Percentage of top 10% of cost cases correctly predicted in top 10%	23.8%	35.4%	26.8%	27.7%
STD of relative weights	0.414	0.564	0.242	0.345
Validation Sample				
R-squared	0.185	0.266	0.127	0.193
Percentage of top 10% of cost cases correctly predicted in top 10%	24.2%	39.7%	27.9%	29.3%
STD of relative weights	0.407	0.561	0.240	0.341
Facility Level R-squared	0.196	0.230	0.066	0.130
Facility-level Cost Models *(Validation Sample)*				
CMI-Only Model				
CMI Coefficient	1.40	1.30	1.54	1.20
R-squared	0.208	0.241	0.085	0.139
Standard Error	0.097	0.094	0.208	0.110
Coefficient different from 1	<0.001	0.002	0.009	0.073
Payment Model				
CMI Coefficient	1.40	1.28	1.52	1.20
R-squared	0.209	0.250	0.096	0.149
Standard Error	0.097	0.094	0.204	0.107
Coefficient different from 1	<0.001	0.003	0.010	0.061
Fully-Specified Model				
CMI Coefficient	1.37	1.34	1.54	1.20
R-squared	0.227	0.258	0.121	0.171
Standard Error	0.099	0.101	0.228	0.120
Coefficient different from 1	<0.001	<0.001	0.019	0.103

Note: Sample sizes for the 10% random sample models are: Acute n=27,105, Chronic n=24,084, Rehabilitation n=110,978, Total sample n=162,167. Sample sizes for the validation sample models are: Acute n=27,903, Chronic n=24,614, Rehabilitation n=116,829, Total sample n=169,346. Sample sizes for all facility level models are: Acute n=1,367, Chronic n=1,347, Rehabilitation n=1,385, Total sample n=1,386 facilities. The payment model includes urban/rural status and area wage index, in addition to the CMI. The fully-specified model includes urban/rural status, area wage index, SNF bed size indicators, hospital-based status, ownership type, and percent Medicare days, in addition to the CMI. Standard errors and p-values reported are the Huber-White standard errors, correcting for heteroskedasticity.

These results suggest that facility-level therapy costs are highly idiosyncratic, with therapy provision apparently being driven by the incentives in the current payment system (providing therapy at the levels that are most profitable), reducing person-level variation based on need and increasing the tendency for use to be driven by facility characteristics and other factors.

The increase in R-squared when facility characteristics are added to the CMI model provides further indication of this. In addition, when we estimated CMI-only models for PT/OT costs in the rehabilitation group separately for nonprofit and for-profit facilities, the results were R-squared values of 0.127 and 0.067 for nonprofit and for-profit homes, respectively, illustrating that the variation in costs is better explained by patient characteristics alone in nonprofit homes than it is in for-profit homes. This suggests that for-profit facilities may use a more standardized, less patient driven, approach to therapy provision. For-profit freestanding SNFs have been shown to have dramatically changed the therapy services they provided following PPS implementation, using largely moderate to high payment categories (White 2003). The current system, therefore, makes it very difficult to estimate models predicting therapy use at the facility level. Overall, the CMI coefficients are greater than 1 and the differences are statistically significant at the 5% level for the individual NP groups but not for the total sample. For example, the payment model for the total sample has a CMI coefficient of 1.2, which was not significantly different from 1.0 at the 5% level but was significant at the 10% level. The R-squared for this model and sample was about 15%. For all facility-level models, the R-squared for the rehabilitation group, where there is least variation in the dependent variable for estimating the prediction model, was the lowest.

Table V.28 shows model fit for wage-adjusted ST costs. In this case, the rehabilitation group had the highest R-squared. The stay-level R-squared values were reasonably well validated in the validation sample, despite the relatively small number of cases receiving speech therapy, and were also comparable at the facility level. Over 40% of the cases in the top 10% in speech therapy were predicted to be in the top 10%, reflecting model strength in terms of the ability to predict cases using substantial ST. The payment model had a CMI coefficient of 1.01 and an R-squared of greater than 20% for the total sample. For all facility-level cost models (total sample), the ST models were more predictive of facility-level costs and better able to capture high-cost patients than the PT/OT models. Over the total samples, the ST models also produced payment weights that would pay facilities in proportion to their ST costs.

Table V.28: Performance Summary for Stay and Facility-Level ST Cost Models

	Acute	Chronic	Rehabilitation	Total sample
Stay-level models				
Test Sample	0.185	0.120	0.207	0.191
R-squared				
Percentage of top 10% of cost cases correctly predicted in top 10%	36.8%	34.0%	45.6%	42.4%
STD of relative weights	1.090	1.036	1.419	1.315
Validation Sample				
R-squared	0.166	0.108	0.191	0.176
Percentage of top 10% of cost cases correctly predicted in top 10%	36.8%	33.1%	46.2%	42.6%
STD of relative weights	1.078	1.045	1.440	1.320
Facility Level R-squared	0.236	0.137	0.120	0.197
Facility-level Cost Models *(Validation Sample)*				
CMI-Only Model				
CMI Coefficient	0.89	0.72	0.91	1.02
R-squared	0.233	0.201	0.197	0.195
Standard Error	0.056	0.035	0.068	0.053
Coefficient different from 1	0.045	<0.001	0.192	0.675
Payment Model				
CMI Coefficient	0.88	0.70	0.90	1.01
R-squared	0.237	0.211	0.205	0.203
Standard Error	0.057	0.036	0.068	0.053
Coefficient different from 1	0.031	<0.001	.153	0.793
Fully-Specified Model				
CMI Coefficient	0.85	0.67	0.87	0.97
R-squared	0.246	0.218	0.214	0.210
Standard Error	0.058	0.038	0.069	0.058
Coefficient different from 1	0.012	<0.001	0.064	0.550

Note: Sample sizes for the test sample models are: Acute n=27,099, Chronic n=24,060, Rehabilitation n=110,956, Total sample n=162,115. Sample sizes for the validation sample models are: Acute n=27,889, Chronic n=24,582, Rehabilitation n=116,810, Total sample n=169,281. Sample sizes for all facility level models are: Acute n=1,367, Chronic n=1,347, Rehabilitation n=1,385, Total sample n=1,386 facilities. The payment model includes urban/rural status and area wage index, in addition to the CMI. The fully-specified model includes urban/rural status, area wage index, SNF bed size indicators, hospital-based status, ownership type, and percent Medicare days, in addition to the CMI. Standard errors and p-values reported are the Huber-White standard errors, correcting for heteroskedasticity.

Discussion

Prior research and the findings presented here suggest that therapy provision in the current RUG-III-based SNF PPS is driven largely by the payment system incentives. The incentive to provide therapy because classification is based on therapy minutes resulted in 85% of SNF patients receiving PT or OT services. Prior work based on the PPS demonstration showed that therapy provision increased in response to implementing the SNF PPS with no accompanying change in outcomes, such as community discharge (Hutt, et al. 2001). Although this may have been an intended consequence of the payment system at the time it was initially developed, because of concerns that therapy services would be underprovided in a PPS unless a fee schedule-like structure was used, the long-term consequences of this incentive are now open to question. A second result of the current payment system is that, because certain payment groups are more profitable than others due to the 20% add-on payments, the amount of therapy provided varies very little across patients. Amount of therapy is generally based on the minimum required to meet the criteria for specific payment categories that are most profitable, resulting in an extremely low coefficient of variation (0.69) for patients who received the majority of the therapy services.

The result of such high use rates and low variation in amount of therapy provided is that the therapy cost dependent variable is problematic for modeling a new payment system. First, with such low variation and high use rates it is difficult to estimate a model that predicts variation at the person level. More fundamental, however, is the fact that the current therapy costs do not necessarily reflect patient therapy needs, so associations between patient characteristics and therapy costs are weakened.

Despite these limitations, we were able to draw some important conclusions about the relationship between patient characteristics and therapy costs. First, the "upside down U" or bow-shaped relationship between function and therapy use was confirmed. Using the Barthel Index, it was evident that patients with the greatest levels of disability and those with the greatest levels of independence used the least amount of therapy based on cost and charges during the hospital stay and SNF stay. This non-linear relationship, which has been demonstrated previously (Eilertsen, et al. 1998), was used to define the rehabilitation patient classification category (i.e., the NP rehabilitation group). That is, those patients in the mid-range, who are most similar to patients in rehabilitation hospitals as well, were placed in the rehabilitation category. In addition, within the NP groups, individual ADLs were found to be associated with therapy use in much the same non-linear manner. Because of this finding, for modeling purposes the ADLs were all split into 3 levels representing independence, moderate dependence, and severe/total dependence rather than treated as a scale.

A second major finding was the importance of separating PT/OT services from ST services in understanding the factors that explain therapy costs. Although ST costs average less than 10% of the cost for PT/OT on average, ST costs were provided to only about 20% of the cases and, when provided, the costs were more comparable to those of PT/OT. In addition to the difference in distribution of ST costs relative to PT/OT, the pattern of associations between patient characteristics and costs is entirely different for PT/OT and ST. In particular, the most severe functionally and cognitively impaired patients used less PT/OT and more ST. Patients with surgical wounds and musculoskeletal diagnoses used more PT/OT and less ST. Certain conditions like swallowing problems were associated with substantially higher ST costs and lower PT/OT costs. These differing associations would not be as evident in models that combine PT, OT, and ST into a single dependent variable.

Third, the patient models estimated separately for PT/OT and ST stratified by the acute, chronic, and rehabilitation categories were reasonably predictive and consistent with our expectations. The PT/OT R-squared was about 19% across all patients in the three groups, with the lowest R-squared in the rehabilitation group where variation was low. The ST person-level R-squared was also 19%. The models also did a reasonable job at predicting high-cost cases, particularly for ST, where 42% of the cases in the top 10% were predicted in the top 10%.

The reduction in R-squared at the facility level for PT/OT requires further investigation. Because this occurred largely in the rehabilitation group, where there was such low variation at the person level, we attribute this largely to low variation in both the predicted value for PT/OT and the actual values for PT/OT at the facility level. It appears that in the current payment environment, variation at the facility level is driven more by factors other than patient characteristics, which are the basis for our facility-level predictive model. When some of the facility predictors are included in the model, the R-squared improves, but not above 17%. However, many have argued that extant variation in therapy costs is neither systematic nor attributable to specific facility characteristics; rather it is associated with such issues as geographic practice patterns and individual facility variation. Several studies have shown geographic variation in use of rehabilitation services (Schlenker, et al. 1997; Lee, et al. 1997). This finding, in fact, provides support for moving towards a payment system for therapy that is based on patient characteristics that are appropriately associated with need for therapy, so that over time therapy is provided more on the basis of need rather than geographic and facility-specific practice patterns. The facility R-squared for ST was more stable, probably reflecting the fact that ST is a relatively scarce resource and provision is probably tied more closely to actual patient need. The CMI coefficients in the ST model were equal to 1.0 and therefore acceptable. The CMI coefficients in the PT/OT model were not significantly different from 1.0, but the point estimate of 1.2 for the total sample is consistent with some (likely manageable) degree of compression, but the estimate is imprecisely measured.

In conclusion, NP-therapy models present a distinct alternative to the existing SNF PPS approach for paying for rehabilitation services, which is with a fee schedule. Recognizing that the current dependent variable of therapy costs does not necessarily reflect actual patient needs for therapy, basing payment on case mix rather than on use should ultimately drive the system towards use patterns based more on patient need.

Two areas are suggested for further development. First, therapy charges from the hospital are included as an independent variable in the current models. Although inflation of therapy charges in response to such a system would not be easy for hospitals to implement and could be easily regulated, alternatives exist. For example, the independent variable could ultimately be an indicator of therapy units of service in the hospital rather than charges, which hospitals would be highly unlikely to game because of the additional costs they would incur by increasing therapy in order to place a patient in an SNF. Even if they did in fact increase the amount of PT/OT provided in the hospital, this may not be a bad result given the evidence that early initiation of therapy is associated with enhanced rehabilitation outcomes.

Second, models incorporating ST with PT/OT to create a single therapy payment classification should be tested. One approach would be to sum the ST predicted costs and the PT/OT predicted costs. Another would be to develop a combined model that includes an indicator of ST use interacted with variables that are correlated in one direction with ST and a different direction with PT/OT. Basically, having an ST indicator in the payment system would provide an explicit incentive for the provision of ST services.

VI. DRG-Based Approaches to Classifying Medicare SNF Patients

A. Aims and Background

Diagnostic Related Groups (DRGs) have been used in the Medicare acute care hospital PPS since 1984; more recently they were adopted to form the basis of the Medicare long-term care hospital PPS. One of the benefits of using DRGs in post-acute care case-mix adjustment is that they are well known and have a history of use. Although DRGs were not designed to explain SNF costs, they can be used to group patients and construct SNF relative weights in order to investigate the extent to which they do predict SNF costs. DRG-based SNF relative weights can be constructed using costs, charges, or some other measure of SNF resource use. DRGs could also be augmented with other variables, such as those identified in the SIM model for NTA and the "New Profiles" models for NTA and therapy.

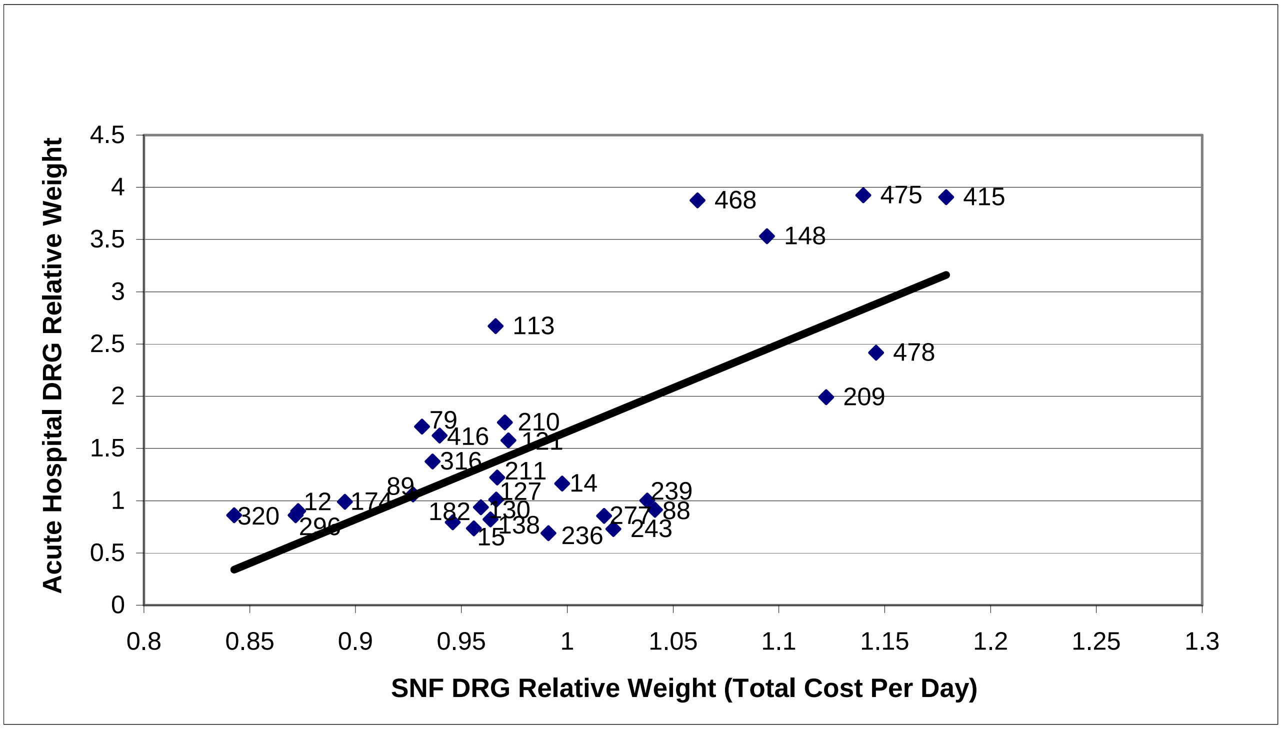
All Medicare SNF patients have a DRG associated with them from the hospital stay preceding their admission to the SNF, since Medicare-eligible SNF stays must be preceded by a minimum three-day hospital stay. Thus, it would be feasible from a data-availability perspective to use the DRG from the qualifying hospital stay in SNF case-mix adjustment. Members from the first TAP generally agreed with the notion that data on diagnosis and conditions from the prior hospital stay are usable in case-mix adjustment for the Medicare SNF PPS. Several concerns have been raised, however, about the administrative effort needed to transmit the hospital DRG code from the hospital to the SNF in a timely manner.

DRGs were designed to group patients with similar clinical problems that are expected to require similar amounts of resources in short-term acute care hospitals. DRGs summarize patient diagnostic information and procedures performed within the hospital, and are further divided by the presence or absence of comorbidities or complications indicated by specific secondary diagnoses (MedPAC 2002). Unlike RUGs, which in part depend on the amount of therapy services a patient receives or is expected to receive, DRGs from the qualifying hospital stay do not depend on the level of care actually provided in the SNF. This is a potentially attractive feature of DRGs, since PPSs typically seek to induce providers to use resources more efficiently by separating payment from resource use, given patients' clinical characteristics. A large majority of members from the first TAP agreed that revised methods should be sought to base rehabilitation therapy payment on patients' clinical characteristics rather than on minutes of therapy provided, as in the existing SNF PPS.

Cotterill (1986) examined the feasibility of using DRGs for Medicare SNF patient classification. They were found to have little explanatory power for routine costs at the facility level but showed greater potential in explaining ancillary costs. To this end, DRGs may hold more promise now than they did in 1986 because ancillary costs now constitute a larger share of SNF costs. If drug and other non-therapy ancillary (ONTA) costs in SNFs are in fact closely related to patient diagnoses, DRGs might be more effective now than before in explaining SNF costs.

In exploratory analyses using SNF cost reports and MDS data from 1999, we constructed DRG relative cost weights for total SNF costs per day. We compared the SNF-DRG relative weights to DRG relative weights used in the acute care hospital PPS. The comparison for the top thirty prior hospital DRGs among Medicare SNF patients is shown in Figure VI.1. The positive relationship between relative SNF costs per day and acute hospital stay relative costs provides *prima facie* evidence of the potential for using DRGs for SNF case-mix adjustment.

Figure VI.1: Total SNF Cost Per Day Relative Weight vs. Acute Hospital Relative Weight (Top 30 DRGs in SNFs 1999)



We also examined the relationship between facility-level SNF costs (routine, therapy, NTA, and total) and a DRG-based facility-level case-mix index (CMI) using the 1999 data. Facility-level cost regressions suggest that DRGs of the prior hospital stay explain a meaningful fraction of variation in SNF costs. For example, the percentage of variation in NTA costs explained increased from 33% (without the DRG-based CMI) to 38% (with the CMI) in an "expanded" model with several other control variables. A similar model for therapy costs showed an increase in R-squared from 8% without the CMI to 19% with it. The model for total costs, however, only showed an increase in R-squared from 57% without the CMI to 60% with the CMI. Most of these models highlighted a serious problem of CMI compression at the facility level when using DRGs by themselves, with CMI coefficients ranging from 1.5 to 2.6. Because limited variance-explanation ability of a classification system is one of the possible causes of CMI compression, these initial findings pointed to the importance of broadening the DRG-based system to incorporate additional patient clinical information, such as comorbidities, and cognitive and physical functioning.

In the February 2004 report to CMS, we presented the results of stay-level models that supplemented the DRGs with functioning and cognitive variables; these variables substantially increased the explanatory power of the cost models. The May 2004 report also considered the use of all patient refined DRGs (APR-DRGs). It found that while APR-DRGs did have some additional predictive ability compared to DRGs, the gain was not large—especially in light of the additional complexity using APR-DRGs would entail. For this report, we limit our consideration of DRG-related approaches to models using DRGs and the functioning variables we identified as the preferred model in the prior report.

In this chapter we more fully document the predictive power of DRGs and the limited set of physical and cognitive functioning variables from the MDS in explaining SNF costs at the person level. We also examine the predictability of the DRG + functioning models at the facility-level. The facility-level models provide evidence on the implications of basing payment on this DRG-based patient classification system for SNFs and for patients. This analysis does not consider the use of specific medical services provided in the SNF, or hospital services or charges from the qualifying hospital stay services (beyond what is captured in the DRGs), for case-mix adjustment. The possible role of such predictors is explored in other chapters of the report. If certain service use variables are deemed useful and appropriate for patient classification, they could be incorporated within a DRG-based system or vice-versa.

B. Data and Methods

The stay-level analyses use test sample for estimation and the stays from the validation sample for validation, as described in Chapter III, with DRGs from the qualifying hospital stay linked to each SNF patient. The test sample contains 163,738 observations for a core set of variables after applying a common set of exclusions. For the purpose of this chapter additional cases were excluded if (1) they did not have a DRG from the qualifying hospital stay, or (2) they had incomplete RUG or cognitive performance scale (CPS) information across all MDS segments within a stay. The resulting restricted test sample used for estimation contains 156,225 cases, and the restricted validation sample contains 169,029 cases.

In analyses of 1999 data, we used DRG grouper software to create DRGs based on diagnoses reported by the SNFs. We found, however, that a very large percentage of hospital-based SNFs used "rehabilitation" as the primary diagnosis, suggesting a systematic difference in coding practices between free-standing and hospital-based SNFs. As a result, the rehabilitation DRG effectively became a hospital-based facility indicator, making the use of DRGs based on SNF diagnoses problematic at the time. Nonetheless, it is still important to note the potential for using diagnostic information provided by the SNF to assign SNF patients to DRGs.

We focus here on the following components of per-day SNF costs: ancillary costs, NTA costs, therapy costs, and total costs. Ancillary costs are measured as the sum of therapy and NTA costs, and total costs are calculated as the sum of ancillary and routine costs. Note that the routine cost component is measured as a constant for all patients within a given facility. This has the effect of muting the overall person-level variation in total costs per day. We also examine ancillary charges, therapy charges, and NTA charges—all per day.

DRG of Qualifying Hospital Stay of SNF Patients

As of March 2002, there were 506 DRGs used for Medicare payment derived from 25 major diagnostic categories. In the estimation and validation samples combined, there are 331 DRGs from the qualifying hospital stay represented among SNF patients. Low-volume DRGs, defined here as DRGs with fewer than 200 SNF patients in the pooled sample, are collapsed into 10 groups based on the case-weighted deciles of average total cost per day for each low-volume DRG. Only 4% of the SNF cases are in low-volume DRGs in both samples. As a result, there are 181 DRG-based groups used in the analyses below, consisting of 171 high-volume DRGs and 10 DRG-based SNF cost groups.

Creating a Functioning Model for Use with DRGs

We sought a parsimonious set of functioning variables to use in conjunction with DRGs to create SNF patient case-mix groups. Candidate functioning variables included individual activities of daily living (ADLs) reported on the MDS, the Barthel index (described earlier in the report), the sum of ADLs, and the cognitive performance scale (CPS). In examining individual ADLs and creating the ADL sum, we treated values of "eight" (activity did not occur) as "four" (fully dependent in activity), just as in the previous chapter on NP approaches. Starting with these indices, we found that the CPS had more predictive power by itself than the Barthel index or the sum of ADLs, so we started with the CPS. We ran regressions of the seven levels of CPS scores (0-6) on the components of costs per day and found that average costs were similar within the following three groups: CPS=0, CPS=1 or 2, and CPS>=3. Notably, the condition CPS>=3 is used in constructing the current RUGs.

We found that individual ADL variables had more predictive power than the indices based on ADLs. This stems from the fact that, for some individual ADLs but not others, a higher level of functional dependence is associated with lower rather than higher SNF costs per day. We ran cost per day regressions using each ADL (expressed as level

1-4) individually as explanatory variables and also ran models with all ADL variables included. From these models, which included the CPS categories we had already identified, three ADL variables stood out as particularly important in explaining costs. They were locomotion outside of unit, locomotion within unit, and eating. Locomotion is widely used in surveys designed to measure health status of the elderly. The eating ADL factors into the assignment of RUGs (although in a different form than we use here) and the Katz ADL index. Thus, these measures have been identified as being clinically meaningful in describing patient functioning in a variety of contexts.

When we examined the ways these variables were related to costs, we found that (1) the top category of the eating ADL (eating dependence) stood out as being more costly than the other categories, and (2) expected therapy and total costs increased monotonically with the locomotion off unit ADL, with categories ADL = 0, ADL = 1, 2, or 3, and ADL = 4 or 8 capturing most of the variance. Locomotion off unit and locomotion within the unit operated similarly, with the former having a stronger relationship to costs. For this reason and because the two locomotion measures were highly correlated, we used only locomotion off unit in our analysis. To summarize, we use a relatively simple framework to describe patient functioning with three CPS categories, three locomotion ADL categories, and two eating ADL categories. The aim of this framework is to add a functional component to a SNF patient classification system based on DRGs.

Stay-Level Regression Models

Regression models were estimated for each cost or charge component, with the dependent variables measured in dollars in our base models. For every cost component, we estimated models using DRGs and the functioning measures as explanatory variables. For each estimated model, as with prior analyses in this report, we present three within-sample measures of predictive ability: R-squared, sensitivity, and the standard deviation of the relative weights. The interpretation and limitations of these measures are discussed in detail in Chapter III. Each of these statistics is also reported for the validation sample, with the coefficients from the test sample used to generate out-of-sample predictions in the validation sample.

SNF costs per day are typically positively skewed. Taking the log of the dependent variables (after adding $1 to deal with individuals whose costs are $0 for a particular component) reduces the skew. In earlier analyses we compare our dollar-level models with log-dollar models, in order to determine whether the predictive performance of the different classifications is sensitive to the choice of a level or log specification. Since the findings suggested that logging the dependent variables is not critical—and may in fact reduce the ability of the models to correctly identify high-cost cases—we present only dollar level models here. Standard errors for stay-level models were estimated to adjust for facility-level clustering (i.e., potential correlations in the models' residuals across patients from the same facility).

It is useful to note how this analysis relates to the discussion of MDS segments versus stays in Chapter III. The analysis here is conducted at the stay level. The dependent variables reflect the covered day-weighted average of patient costs per day over the stay. The DRG measures are constant over the course of the stay, being determined by information from the qualifying hospital stay. The CPS score we use is the score obtained from elements of the five day MDS. In principal, the CPS score could change over the course of the stay, in which case using the measure from only the first MDS would not make full use of the data. In practice, however, the CPS score was found to be constant over a stay for the vast majority of patients. This constancy is not surprising, particularly for Medicare stays, which average only twenty-three days. ADL scores, in contrast, vary substantially over multi-segment stays. The analyses reported in this chapter incorporate this variation. Specifically, the ADL variables are defined as the percent of MDS segment covered days during the stay with the given ADL. Under a payment system based on these categories, payments could change if functioning changes during the SNF stay (as is now the case with the RUGs), but the DRG would be fixed.

Facility-Level Regression Models

After generating predicted costs or charges—based on the DRG + functioning classification system, at the stay-level within the validation sample—we compute CMIs for each facility represented in the validation sample. Recall that the validation sample contains all stays within a given facility. As described in Chapter III, the log of the CMI is then used as an explanatory variable in a regression of the log of facility-level costs, with separate regressions run for each cost component with its associated CMI. The log-log specification is used so that the coefficient on the CMI variable is an elasticity, which measures whether or not payments are proportional to costs at the facility level. This formulation provides a test for CMI compression. We report the facility-level R-squared, coefficient on the (log of the) CMI, the standard error of the CMI coefficient, and the p-value for the two-sided test of the null hypothesis that the CMI coefficient is equal to 1.

We estimate the three specifications described in Chapter III: a CMI only model, a payment model, and the fully-specified model. We test for CMI compression and discuss its implications for each specification. These implications are described in Chapter III.

C. Findings

Summary of Cost and Charge Measures

Wage-adjusted cost and charge measures are summarized in Table VI.1. The average total cost per day for the test sample is $339, with a standard deviation (SD) of $180. Ancillary costs average $135 per day, with a coefficient of variation (CV) of 0.71, even before making any adjustments for case-mix. A rule of thumb for patient classification systems is that the case-mix groups should have a CV below 1.0. Ancillary costs pass this criterion for the grand average, even before grouping patients. This suggests that, relative to other PPS settings, there is less absolute variation to start with in patient ancillary cost per day (which the patient classification system needs to reduce).

Table VI.1: Summary of Wage-Adjusted SNF Cost and Charge Measures

	Mean	SD	CV
Costs per day			
Ancillary	$135	$96	0.71
NTA	$63	$82	1.30
Therapy	$73	$55	0.75
Total	$339	$180	0.53
Charges per day			
Ancillary	$271	$304	1.12
NTA	$144	$273	1.90
Therapy	$127	$105	0.83

Note: Test sample. N=156,225. SD = standard deviation. CV = standard deviation of cost (charge) divided by the average cost (charge). Means for the validation sample (not reported) are nearly identical.

Table VI.1 also illustrates that charges vary more than costs. On the one hand, differences in the ways facilities report charges may induce artificial variation relative to costs or actual resource use. On the other hand, costs assume constant cost-to-charge ratios (CCRs) for each cost component, which may artificially reduce variation in our cost measures relative to actual resource use. Because of limitations in both cost and charge data, neither measure can be argued to be superior to the other on conceptual grounds.[23] For this reason, we investigate the difference in variation between cost and charge relative weights. We also examine the correlation between the alternative sets of weights to evaluate whether relative payments would be similar for similar patients under the two sets.

[23] Rogowski and Byrne (1990) make this point on the merits of cost versus charge relative weights in the context of the acute hospital PPS.

Description of DRGs in SNFs

Table VI.2 lists the thirty most common DRGs, their descriptors, their means and standard deviations, and their coefficients of variation for total SNF costs per day. The most common DRG, 209 – Major Joint and Limb Reattachment Procedures—covers nearly 8% of SNF patients and costs $397 on average. Of the thirty most common DRGs, the least costly is DRG, 429 – Organic Disturbances and Mental Retardation—which costs $239 on average. Table VI.3 lists the most costly DRGs in SNFs with at least 100 cases in the test sample. The most costly DRG, 471—Bilateral or Multiple Joint Procedure—costs $452 per day and does not appear on the list of most common DRGs. The most common DRG, 209, happens to be the 12th most expensive (at $397). The only other DRG on both lists is 478 – Other Vascular Procedures w/CC—25th on the most common list and twenty-second on the most costly, at an average cost of $379 per day.

Table VI.2: 30 Most Common DRGs in SNFs

DRG	Description	N	Mean	SD
209	MAJOR JOINT & LIMB REATTACHMENT PROCEDURES OF LOWER EXTREMITY	12,371	$397	$197
14	SPECIFIC CEREBROVASCULAR DISORDERS EXCEPT TIA	8,261	325	168
89	SIMPLE PNEUMONIA & PLEURISY AGE >17 W CC	8,188	308	182
127	HEART FAILURE & SHOCK	7,814	321	178
210	HIP & FEMUR PROCEDURES EXCEPT MAJOR JOINT AGE >17 W CC	7,385	322	172
296	NUTRITIONAL & MISC METABOLIC DISORDERS AGE >17 W CC	4,826	288	159
320	KIDNEY & URINARY TRACT INFECTIONS AGE >17 W CC	4,231	275	149
79	RESPIRATORY INFECTIONS & INFLAMMATIONS AGE >17 W CC	4,088	316	197
416	SEPTICEMIA AGE >17	3,555	310	187
88	CHRONIC OBSTRUCTIVE PULMONARY DISEASE	3,471	342	196
462	REHABILITATION	2,913	298	132
121	CIRCULATORY DISORDERS W AMI & MAJOR COMP, DISCHARGED ALIVE	2,875	313	171
174	G.I. HEMORRHAGE W CC	2,804	290	161
148	MAJOR SMALL & LARGE BOWEL PROCEDURES W CC	2,492	362	198
243	MEDICAL BACK PROBLEMS	2,241	325	161
12	DEGENERATIVE NERVOUS SYSTEM DISORDERS	2,100	263	134
236	FRACTURES OF HIP & PELVIS	2,019	321	163
182	ESOPHAGITIS, GASTROENT & MISC DIGEST DISORDERS AGE >17 W CC	1,954	306	170
316	RENAL FAILURE	1,855	310	182
475	RESPIRATORY SYSTEM DIAGNOSIS WITH VENTILATOR SUPPORT	1,695	365	201
429	ORGANIC DISTURBANCES & MENTAL RETARDATION	1,689	239	110
211	HIP & FEMUR PROCEDURES EXCEPT MAJOR JOINT AGE >17 W/O CC	1,646	334	180
239	PATHOLOGICAL FRACTURES & MUSCULOSKELETAL & CONN TISS MALIGNANCY	1,627	335	170
277	CELLULITIS AGE >17 W CC	1,598	344	197
478	OTHER VASCULAR PROCEDURES W CC	1,549	379	200
138	CARDIAC ARRHYTHMIA & CONDUCTION DISORDERS W CC	1,464	318	171
113	AMPUTATION FOR CIRC SYSTEM DISORDERS EXCEPT UPPER LIMB & TOE	1,424	318	184
15	TRANSIENT ISCHEMIC ATTACK & PRECEREBRAL OCCLUSIONS	1,419	308	157
468	EXTENSIVE O.R. PROCEDURE UNRELATED TO PRINCIPAL DIAGNOSIS	1,356	341	186
130	PERIPHERAL VASCULAR DISORDERS W CC	1,273	308	172

Table VI.3: 30 Most Costly DRGs in SNFs (With at Least 100 Cases in the Test Sample)

DRG	Description	N	Mean	SD
471	BILATERAL OR MULTIPLE MAJOR JOINT PROCS OF LOWER EXTREMITY	339	$452	$181
126	ACUTE & SUBACUTE ENDOCARDITIS	144	429	213
105	CARDIAC VALVE & OTHER MAJOR CARDIOTHORACIC PROC W/O CARDIAC CATH	370	423	193
104	CARDIAC VALVE & OTHER MAJOR CARDIOTHORACIC PROC W CARDIAC CATH	385	420	206
498	SPINAL FUSION EXCEPT CERVICAL W/O CC	177	419	187
109	CORONARY BYPASS W/O CARDIAC CATH	500	417	201
497	SPINAL FUSION EXCEPT CERVICAL W CC	421	410	201
483	TRACHEOSTOMY EXCEPT FOR FACE,MOUTH & NECK DIAGNOSES	760	409	210
191	PANCREAS, LIVER & SHUNT PROCEDURES W CC	109	404	196
499	BACK & NECK PROCEDURES EXCEPT SPINAL	385	403	199
107	CORONARY BYPASS W CARDIAC CATH	919	400	195
209	MAJOR JOINT & LIMB REATTACHMENT PROCEDURES OF LOWER EXTREMITY	12,371	397	197
479	OTHER VASCULAR PROCEDURES W/O CC	128	396	201
500	BACK & NECK PROCEDURES EXCEPT SPINAL FUSION W/O CC	223	395	187
20	NERVOUS SYSTEM INFECTION EXCEPT VIRAL MENINGITIS	120	394	204
75	MAJOR CHEST PROCEDURES	325	394	203
452	COMPLICATIONS OF TREATMENT W CC	188	391	211
216	BIOPSIES OF MUSCULOSKELETAL SYSTEM & CONNECTIVE TISSUE	201	390	188
231	LOCAL EXCISION & REMOVAL OF INT FIX DEVICES EXCEPT HIP & FEMUR	225	386	193
304	KIDNEY,URETER & MAJOR BLADDER PROC FOR NON-NEOPL W CC	117	384	227
154	STOMACH, ESOPHAGEAL & DUODENAL PROCEDURES AGE >17 W CC	508	380	209
478	OTHER VASCULAR PROCEDURES W CC	1,549	379	200
292	OTHER ENDOCRINE, NUTRIT & METAB O.R. PROC W CC	114	378	184
110	MAJOR CARDIOVASCULAR PROCEDURES W CC	559	376	197
418	POSTOPERATIVE & POST-TRAUMATIC INFECTIONS	319	372	186
217	WND DEBRID & SKN GRFT EXCEPT HAND,FOR MUSCSKELET & CONN TISS DIS	557	372	194
76	OTHER RESP SYSTEM O.R. PROCEDURES W CC	683	372	209
115	PRM CARD PACEM IMPL W AMI,HRT FAIL OR SHK,OR AICD LEAD OR GNRTR PROC	183	370	184
238	OSTEOMYELITIS	259	370	179
170	OTHER DIGESTIVE SYSTEM O.R. PROCEDURES W CC	181	368	213

Description of Functional Variables and Relationship with SNF Costs

About 27% of SNF patients had a CPS score of 1 or 2 (see Table VI.4). Slightly over one-third (35%) of patients were in the most impaired group, with a CPS of 3 or higher. The remainders were in the reference group (CPS = 0), meaning no cognitive impairment. Slightly over half (54%) of patients (on average over patients and MDS segments) were dependent in locomotion off unit or did not perform the activity. About 13% of patients were in the most dependent group for eating.

Table VI.5 presents the effects of the functioning variables on costs per day in regressions that also control for DRGs. Being more cognitively impaired is associated with lower SNF costs per day for each component of cost. Being more dependent in locomotion off unit is associated with higher ancillary, NTA (for the top category), and total costs per day, but the effects on therapy costs per day are not monotonic. Being dependent in eating is associated with lower therapy costs but higher NTA costs, presumably because people who are unable to eat without a great deal of assistance are unable to engage in therapy but do require additional resources of other types of services (such as intravenous nutrition). Total SNF costs are about $30 lower for the eating dependent group.

Table VI.4: Description of Functional Variables

	Mean
CPS = 0	(Reference group)
CPS = 1 or 2	0.267
CPS ≥ 3	0.350
Locomotion off unit ADL = 0	(Reference group)
Locomotion off unit ADL = 1, 2, or 3	0.381
Locomotion off unit ADL = 4 or 8	0.537
Eating ADL = 0, 1, 2, or 3	(Reference group)
Eating ADL = 4 or 8	0.134

Note: N = 156,225. CPS refers to the Cognitive Performance Scale measured on the 5 day MDS, which is increasing in impairment. Higher numbers for ADLs imply more dependence. ADLs are measured as day-weighted averages over the MDS segments within a stay.

Table VI.5: Cost Per Day Regression Results for Functional Variables

	Ancillary	NTA	Therapy	Total
CPS = 1 or 2	-$14.6	-$12.6	-$1.98	-$58.0
	(0.795)	(0.641)	(0.448)	(2.11)
CPS ≥ 3	-31.4	-18.9	-12.5	-97.7
	(0.886)	(0.712)	(0.523)	(2.32)
Locomotion off unit ADL (1, 2, or 3)	18.5	0.371	18.1	52.6
	(1.19)	(0.900)	(0.749)	(2.98)
Locomotion off unit ADL (4 or 8)	35.4	20.7	14.6	125
	(1.36)	(1.05)	(0.805)	(3.65)
Eating ADL = (4 or 8)	-1.26	21.0	-22.2	-30.0
	(1.19)	(1.02)	(0.574)	(2.42)
R-squared	0.055	0.076	0.134	0.128

Note: N = 156,225. OLS coefficients reported. Standard errors of regression coefficients are reported in parentheses. Standard errors are adjusted for clustering by facility.

Predictive Ability of DRGs and Functioning Variables for SNF Costs at the Stay-Level

The key results of this analysis are presented in Table VI.6, which reports predictive measures for models of each cost component using DRGs + functioning as the explanatory variables. We focus on the results for the validation sample (because we are most interested in out-of-sample performance), noting that results are quite similar for the test sample. While the large number of DRGs may introduce the potential for model overfitting, comparing the estimation and validation sample results suggests overfitting is not a problem in our case. It is normal to find some depletion in R-squared in the validation sample relative to the test sample, but the depletion we observe is minor.

Table VI.6: Models of Person-Level SNF Costs by Component Using DRGs + Functioning as Explanatory Variables

		Ancillary	NTA	Therapy	Total
Test sample					
	R-squared	0.055	0.076	0.134	0.128
	Sensitivity	0.199	0.261	0.242	0.248
	SD of relative weights	0.167	0.359	0.278	0.190
Validation sample					
	R-squared	0.048	0.071	0.125	0.109
	Sensitivity	0.197	0.258	0.244	0.231
	SD of relative weights	0.167	0.360	0.279	0.189

Note: N = 156,225 in the estimation sample and N = 169,029 in the validation sample. Model degrees of freedom = 186 in all models.

DRGs + functioning variables explain 4.8% of ancillary costs, with a sensitivity of 20%. The standard deviation of the relative weights that would result from this patient classification system is about 17%. When we break overall ancillary costs into its two major components, NTA and therapy costs, we find more underlying variation in the relative weights, and higher predictability measures. The R-squared is 7.1% for NTA and 12.5% for therapy. The fact that the dispersion of the relative weights is lower in the total ancillary model merely reflects the fact that on average, higher predicted NTA costs are associated with lower predicted therapy costs. It is important to keep in mind that SNF ancillary payments would be exactly the same if payments were based on separate NTA and therapy models with the payments summed, as they would be if payments were based on the single ancillary model.

Using DRGs +functioning to predict total SNF costs, including the nursing/routine component, we find that the R-squared is about 11% in the validation sample, and sensitivity is 23%. This is substantially more than would be achieved if the forty-four RUG groups were used to predict total costs (see the DRG comparisons in Chapter VII for details) and suggests that the DRG model has substantial ability to predict routine costs. This is remarkable considering that the DRGs were not originally designed to predict SNF resource use, whereas the RUGs were. The functioning variables contribute substantially to this finding, but with an intentionally limited set of variables that includes only three MDS measures.

Predictive Ability of DRGs and Functioning Variables for SNF Charges at the Stay-Level

Analogous results for SNF charges are presented in Table VI.7. The R-squared in the charge models are higher than in the cost models for the ancillary costs (7.0% vs. 4.8%) and the therapy models (14.2% vs. 12.5%), but about the same for the NTA models. Since R-squared values are not directly comparable across models with different dependent variables, R-squared is an unreliable metric for comparing the charge and cost models. The other measures of predictive ability may be preferable in this regard. The charge-based models are generally better able to distinguish the top decile cases. The gain is especially pronounced for ancillary charges relative to ancillary costs.

The standard deviations of the relative weights have absolute interpretations, since they measure the variation in payments SNFs would receive for different types of patients under the different methods for computing the relative weights for a given classification system. We find that the relative weights are more variable for the charge models than the cost models for ancillary (29% vs. 17%), NTA (50% vs. 36%), and to a more limited degree, therapy (31% vs. 28%).

The large differences in the variation of the relative weights in these findings suggest very important payment implications surrounding the decision of whether to base relative weights on costs versus charges. We describe the implications in more detail below, by examining whether the charge-based relative weights rank patients' relative costs differently than cost-based relative weights.

Table VI.7: Models of Person-Level SNF Charges by Component Using DRGs + Functioning as Explanatory Variables

		Ancillary	NTA	Therapy
Test sample				
	R-squared	0.069	0.071	0.138
	Sensitivity	0.253	0.273	0.252
	SD of relative weights	0.294	0.504	0.309
Validation sample				
	R-squared	0.070	0.072	0.142
	Sensitivity	0.246	0.269	0.264
	SD of relative weights	0.293	0.504	0.310

Note: N = 156,225 in the estimation sample and N = 169,029 in the validation sample. Model degrees of freedom = 186 in all model

Correlations between Cost- and Charge-Based Relative Weights

While the relative weights based on charges are more spread out than those based on costs, the groups defined by the DRGs and functioning variables are ranked similarly, as evidenced by the correlations in Table VI.8. The correlation between the cost- and charge-based relative weights is 0.95 for ancillary, 0.98 for therapy, and 0.95 for NTA.

Table VI.8: Correlations Between Cost and Charge Relative Weights

	Ancillary costs	Ancillary charges	Therapy costs	Therapy charges	NTA costs	NTA charges
Ancillary costs	1					
Ancillary charges	**0.95**	1				
Therapy costs	0.51	0.32	1			
Therapy charges	0.62	0.47	**0.98**	1		
NTA costs	0.59	0.71	-0.39	-0.26	1	
NTA charges	0.72	0.86	-0.19	-0.04	**0.95**	1

Note: N = 156,225. Relative weights are computed from models using DRGs and functional measures as explanatory variables.

The correlations between the therapy and NTA cost and charge weights show an interesting pattern. The cost weights are negatively correlated, illustrating that patients predicted to cost more than average for therapy are predicted to cost less than average for NTA services. Similarly, the charge weights for therapy and NTA are negatively correlated, but the correlation coefficient is very small, close to 0. If expected therapy and NTA resource use are truly negatively correlated, the low absolute correlation for charges may be reflective of facility charging practices common to therapy and NTA services. Alternatively, if there is no clinical reason for expected therapy and NTA resource use to be negatively correlated, it is worth further consideration as to whether the way CCRs are computed could induce an artificial negative correlation between the therapy and NTA cost relative weights.

Given that the cost- and charge-based relative weights rank patients similarly, but the charge-based weights imply more dispersion in payments, the critical question becomes: Which set of weights more closely tracks actual relative resource use within SNFs? The facility-level analyses provide some insight on this question.

Facility-Level SNF Cost Models Using Cost-Based CMIs

Results for the facility-level cost regressions using the cost-based CMIs are presented in Table VI.9. The facility-level R-squared for (wage unadjusted) ancillary costs is 7.5%. Including wage index and urban variables increases R-squared only to 8.3%. We find evidence of CMI compression in the payment regression, suggesting that if facilities were paid according to the DRGs + functioning classification system, high CMI facilities would be paid at less than cost for their more-costly case-mix, and low CMI facilities would be paid at more than cost for their less-costly case-mix.

Table VI.9: Models of Facility-Level Costs Using Cost-Based DRGs + Functioning CMIs[a]

Explanatory Variables	Ancillary	NTA	Therapy	Total
CMI only model				
R-squared	0.075	0.136	0.077	0.248
CMI coefficient	1.51	1.46	1.23	2.00
SE on CMI coefficient	0.161	0.117	0.129	0.116
p-value	<0.005	<0.005	0.083	<0.005
Payment model[b]				
R-squared	0.083	0.144	0.092	0.289
CMI coefficient	1.49	1.45	1.29	1.96
SE on CMI coefficient	0.161	0.118	0.130	0.115
p-value	<0.005	<0.005	0.027	<0.005
Fully-specified model[c]				
R-squared	0.180	0.289	0.129	0.577
CMI coefficient	0.461	0.925	1.22	0.467
SE on CMI coefficient	0.190	0.115	0.137	0.096
p-value	<0.005	0.515	0.109	<0.005

a. Costs, CMIs, and the area wage index are entered in log form (see methods section in Chapter III).
b. The payment model includes urban/rural status and the area wage index, in addition to the CMI.
c. The fully-specified model includes urban/rural status, area wage index, indicators for number of SNF beds (0-9 beds, 10-19 beds, 20-49 beds, 50-99 beds, with 100+ beds as the excluded category), hospital based dummy, ownership type (voluntary/nonprofit dummy, for profit other dummy, and government dummy, with for profit chain as the excluded category) and percent Medicare days, in addition to the CMI.

When additional control variables are added to the model, the R-squared for ancillary costs increases to 18%, suggesting the strong influence of facility characteristics net of case-mix. Furthermore, the CMI coefficient becomes significantly less than 1.0 in the fully-specified model, indicating CMI decompression. This suggests that within a facility of a given type, facilities would have incentives to select more costly patients for whom SNFs are relatively overpaid and select against low cost patients for whom they are relatively underpaid. The fact that we see evidence of compression without controls, but decompression with controls, highlights the very strong role of facility characteristics relating to both facility costs and facility CMIs. One possible explanation is that the DRG + functioning classification system is relatively weak, and the predictive ability it does have might be picking up facility cost characteristics to a substantial degree.

For facility-level NTA costs, we find evidence of compression in the payment models but not in the fully-specified model, where the CMI coefficient is 0.93. This suggests there could be some payment inequities across facilities, but only a modest incentive to select against certain types of patients within facilities. We note, however, that the standard errors on the CMI coefficients are large (almost 12% for NTA), implying that our estimates of compression are imprecise.

The therapy models indicate CMI compression in both the payment and fully-specified models. It is worth noting that the R-squared is relatively low in the facility-level models for therapy, even when controls for facility-type are added. This suggests that the extent to which facilities incur therapy costs is highly idiosyncratic across facilities, implying that the amount of therapy a SNF patient receives may be largely luck.

DRGs + functioning explain a relatively large share of variation in total costs (25%) in the CMI only model. This is substantially higher than we find for the RUG-III groups (see the comparisons for DRGs in Chapter VII). But there is a large degree of CMI-compression in the total cost models, with a CMI coefficient of nearly 2.0. In the fully-specified model, R-squared is nearly 58% for total costs, but the CMI coefficient is only 0.47, evidence of decompression. This suggests that basing payments on the relative weights we constructed for the DRG + functioning classification system would create incentives for providers to select patients in favor of those expected to be the most costly. We revisit these issues in Chapter VII, where we compare the results of the DRG-based system for total costs with those of the current SNF PPS.

Facility-Level SNF Cost Models Using Charge-Based CMIs

Results for the facility-level cost regressions using the charge-based CMIs are presented in Table VI.10. Recall that even though the CMIs are charge-based, we use facility-level cost as the dependent variable. This is because costs are the preferred measure of resource use at the facility level (see Chapter III), whereas at the stay-level, it is debatable whether estimated costs or reported charges are better measures of resource use.

Table VI.10: Models of Facility-Level Costs Using Charge-Based DRGs + Functioning CMIs[a]

	Ancillary	NTA	Therapy
CMI only model			
R-squared	0.057	0.121	0.069
CMI coefficient	0.677	0.828	0.989
SE on CMI coefficient	0.086	0.077	0.108
p-value	<0.005	0.026	0.919
Payment model[b]			
R-squared	0.065	0.130	0.082
CMI coefficient	0.661	0.823	1.03
SE on CMI coefficient	0.086	0.077	0.108
p-value	<0.005	0.022	0.804
Fully-specified model[c]			
R-squared	0.176	0.269	0.120
CMI coefficient	0.106	0.441	1.01
SE on CMI coefficient	0.095	0.075	0.120
p-value	<0.005	<0.005	0.962

a. Costs, CMIs, and the area wage index are entered in log form (see methods section in Chapter III).
b. The payment model includes urban/rural status and the area wage index, in addition to the CMI.
c. The fully-specified model includes urban/rural status, area wage index, indicators for number of SNF beds (0-9 beds, 10-19 beds, 20-49 beds, 50-99 beds, with 100+ beds as the excluded category), hospital based dummy, ownership type (voluntary/nonprofit dummy, for profit other dummy, and government dummy, with for profit chain as the excluded category) and percent Medicare days, in addition to the CMI.

The facility-level R-squared values are slightly lower when we use the charge-based rather than the cost-based CMIs. Given that the cost and charge-based relative weights have correlations near 1.0, we should not expect large differences in R-squared. It comes down to a question of scale, which the CMI coefficients help sort out. Across all cost components and models, the charge-based CMIs imply lower CMI coefficients. Using the charge-based CMIs, we find evidence of decompression in both the payment models and the fully-specified models for ancillary costs and NTA costs—with charge-based CMIs exacerbating both the extreme decompression in the fully-specified ancillary model *and* the mild decompression in the fully-specified NTA model.

For the therapy models, in contrast, which exhibited compression using the cost-based CMIs, we find proportionality using the charge-based CMIs in both the payment and fully-specified models. Recall that, in the stay-level analyses, the choice of using cost versus charges to construct relative weight made the smallest difference for the therapy models. The relative weights from the charge-based stay-level models were somewhat more dispersed, however, suggesting that the additional dispersion was just enough to correct the compression problem we see using the cost-based CMIs.

D. Discussion

The DRG approaches to patient classification presented here would be a marked shift from the current system based on RUG-III. One benefit of the DRG approaches is that they are based solely on patient characteristics—unlike the current RUGs, which depend in part on the amount of therapy received or merely anticipated. Functional dependency is considered the most important driver of SNF nursing and therapy resource use, and there have been concerns and statistical evidence that DRGs by themselves do not sufficiently distinguish patients with different levels of functional dependency. For this reason, we have shown in this chapter how DRGs can be enhanced with functional indicators to better explain variation in resource use as measured by costs or charges.

Admittedly, the approach presented here considers only patient classification. But the DRG-based approach can also be considered in combination with other aspects of payment policy, such as geographical or facility adjustments or outlier payments.

Statistical Significance

The amount of variation explained in this approach—combining DRGs and five functional variables—differs depending on whether costs or charges are used as dependent variables. In the validation sample, the approach explains 4.8% of ancillary costs, 7.1% of NTA cost, 12.5% of therapy costs, and 10.9% of total SNF costs. The approach explains 7.0% of ancillary charges, 7.2% of NTA charges, and 14.2% of therapy charges. Because differences in R-squared across models with different dependent variables are not directly comparable, we examined differences in relative weights, which let us directly compare the dispersion of relative payments. Relative weights based on charges vary substantially more than relative weights based on costs in models that use the same explanatory variables. Because both costs and charges can differ from true resource use, it is not obvious which relative weights are closer to the unobserved "true" relative weights. The facility-level regressions are informative in this regard, and suggest that using charge-based relative weights exacerbates problems of decompression in the fully-specified models for ancillary and NTA costs. The general tendency of the charge-based relative weights to be more dispersed is actually beneficial in counteracting the CMI compression observed in the models using cost-based CMIs for therapy costs, producing proportionality between facility-level therapy costs and the charge-based CMI for therapy.

Clinical Meaningfulness

The classification methods based on DRGs described here should be clinically meaningful, since the DRGs are already used in other settings (including one post-acute care setting—long-term care hospitals). Of course, this does not imply that they are necessarily sufficient to describe the clinical needs of SNF patients, which motivated our examination of measures of functional dependence. The CPS and ADLs are already accepted as valid measures of patient functioning and are currently used in various forms within the RUGs. The variant considered here makes an additional cut, based on the CPS score beyond what is used in the RUGs. The eating ADL is used in both the Katz ADL index and the RUGs. More broadly, ADL dependencies were designed to measure the extent of disability; these require types of assistance other than medical care and are generally regarded as clinically meaningful.

Administrative Burden

DRG-based approaches would involve the burden of shifting to a different system in SNFs, but this burden is mitigated by the fact that the DRGs are already well-known and used in other systems. The functional data required for the systems described here are already collected on the MDS, and the DRG information is available from the qualifying hospital stay. One question is whether the DRG information can be transmitted to the SNF in a timely manner for payment purposes. Even if there is some time lag, this would not invalidate the approach, since a provisional DRG code could be used initially, with payments reconciled at a later date once the DRGs become available. Alternatively, although we did not study this possibility for the current report, if diagnostic information was reported and coded more completely and consistently across SNFs, DRG grouper software could be applied to diagnoses recorded by the SNF to obtain DRGs for payment purposes.

Financial Incentives

DRGs are clinically based, in that they were designed to capture differing resource needs of acute-care hospital patients. They have the further advantage that, because they are determined before the SNF stay (i.e., are truly prospective), SNFs are not able to manipulate them. This eliminates possibilities for gaming, at least for the component of payment that is tied to the DRG. Payments can still adjust as SNF patients needs change, however, through the functional component we have introduced into this DRG-based classification system.

Going back a step in the overall Medicare episode, we see that hospital-based SNFs may have some ability and incentive to game the DRGs, since they would be able to optimize payment across the entire Medicare episode from the acute care stay to the SNF. We think it unlikely that hospitals would game the DRG of the acute care hospital stay to increase subsequent payment for SNFs, however, particularly since the increase in SNF payment would be relatively small in most cases. Nonetheless, it is a possibility that should be considered in evaluating the different payment and classification system options.

Conclusion

The DRG-based approach, with the addition of a functional dependency component, offers a workable alternative to the current system, based entirely on patient characteristics that have been found useful in short-term acute hospitals, SNFs, and other settings. In terms of variance explanation, however, DRG approaches will generally be dominated by approaches that include service use measures as predictors. In that case, the relevant question becomes the extent to which any gain in variance explanation comes at the cost of weakening the PPS incentive to provide services efficiently, or the risk of producing incentives to over-utilize services that come with an explicit payment enhancement.

Because we have been intentionally parsimonious in our use of clinical measures—by using a sparse set of functional measures and DRGs without modification (other than collapsing some low-volume groups)—other classification systems explain more variance than the approach used here. A more detailed set of functional variables would likely increase the R-squared beyond what is reported here and could be combined with DRGs.

Using additional variables would likely strengthen the relationship between CMIs and facility-level costs, and also make the relationship more proportional. This is important to note, since the main problem with using the DRG + functioning classification system without further enhancement is the compression and decompression patterns we found in the facility-level analyses. One explicit motivation for not including some of the “kicker” variables used in the

RUG-refinement and NP approaches in our DRG-based approaches is to illustrate how far we can go using only off-the-shelf DRGs and a parsimonious set of functioning variables.

Using an outlier payment in conjunction with the DRG-approach would offer the possibility of addressing some of the limitations of the DRG-approach. This would compensate for any potential shortfall in variance explanation (relative to what would be considered ideal) of DRG-based approaches or other approaches based purely on patient clinical characteristics.

Because the DRG-based option would change the way SNFs are paid, there might be some resistance from SNFs, especially SNFs that provide large amounts of therapy, as payment would no longer be keyed to the quantity of therapy services provided. By explaining NTA costs better than the current RUG-III, however, facilities that provide high levels of NTA services could benefit from being reimbursed more proportionately to the services provided. Basing SNF payments to a degree on DRGs would be a step in the direction of consistency in how post-acute care is financed under Medicare, since DRGs are also used in the long-term care hospital PPS. Bringing more consistency to how Medicare pays for post-acute care could be viewed as beneficial.

VII. Comparison of Models

A. Introduction

In considering policy relevance, it is useful to highlight the research discussed in preceding chapters in terms of the five basic options that could potentially be used to modify the case-mix classification system for Medicare SNF patients. As noted, these options could serve as building blocks to address different components of Medicare SNF payments, either alone or concomitantly.

The first three options attempt to classify patients according to NTA costs, and could effectively expand or supplement the current RUG-III system. The fourth option addresses rehabilitation therapy and could replace the existing component of the RUG-III system by adopting a "need-based" rehabilitation therapy classification model. The fifth option is a DRG-related model originally created to classify acute-care hospital patients. Because the SNF benefit requires a prior stay in such a hospital, we examined DRG-related models as a classification system for total SNF costs, as well as for components such as rehabilitation therapy and NTAs.

In this chapter, we highlight the features of the five classification options. Variants of these models can readily be defined within the framework of each option. We then explore the pros and cons of each, and compare them with the current RUG-III system. In discussing the pros and cons, we examine conventional criteria, such as clinical meaningfulness, statistical significance, provider burden, ease of implementation and administration, and appropriate clinical and financial incentives. We incorporate input from our TAP members on the pros and cons of the options.

To focus our discussion, we will first examine the three models developed to address NTA services. We will then discuss the "need-based" rehabilitation model relative to the "fee schedule-like" RUG-III rehabilitation therapy component. Finally, we will address the DRG-related model for total cost, and compare findings with corresponding results for total costs and total payments under the current RUG-III system.

B. NTA-Specific Options

All three of our NTA-specific options were designed to improve the classification of SNF patients with regard to NTA costs. The three options can be viewed as a continuum of choices based on model complexity. The three models are: (a) RUG-58, (b) the Service Index Model+RUG58G (SIM), and (c) our New Profile-NTA (NP-NTA).

These options aim to address patient variation in NTA costs. Adaptation of any of these models is not intended to undo the existing RUG-III classification system's use in adjusting payments for nursing services and rehabilitation therapy services. In effect, payments for NTA services, currently embedded in the nursing payment component, would be carved out and distributed according to one of the NTA-specific options.

Models

RUG-58

This method of classifying patients according to NTA costs was developed in prior research (Abt Associates 2000). It entails creating fourteen additional categories in the RUG-III classification system to account for patients who qualify for both the “extensive services” and “rehabilitation” categories of that system. Prior research (Abt Associates 2000) showed that such patients have much higher NTA costs than patients who qualified for only the rehabilitation category. As such, these fourteen additional groups are simply added to the top of the RUG-III hierarchy (see Chapter IV).

Pros:

- Implementation and administration requirements are minimal
- Providers and policy communities are knowledgeable about this option
- It uses the same data items currently required for the RUG-III system but is a better predictor of NTA costs
- It is based solely on SNF MDS data

Cons:

- It has a relatively low variance explanation of NTA costs (10%)
- It does not capture key subgroups of patients with high NTA costs

Service Index Model (SIM)

This option extends the RUG-58 model and entails two basic steps (see Chapter IV). First, after patients are assigned to the RUG-58 classification system, they are grouped into eight broad categories according to the first digit of the RUG-58 category. For example, the first two categories are: (1) those who qualify for both extensive services and rehabilitation, and (2) those who qualify only for rehabilitation. Second, the eight-category RUG-58 variable is included with five other MDS and claims-based patient condition and service use variables, to assign patients, based on statistical analyses, to NTA payment categories.

Two of the more important variables—IV medication and respiratory therapy (RT)—in the SIM model are derived from the interaction of indicators of service use from the MDS assessments with those from SNF claims. This construct ensures that the services were provided during the SNF stay, since indicators of service use from only the MDS can also refer to prior hospital stays. Although it is possible to use this method to operationalize the variables under a payment system, it would be cumbersome. It would be considerably more desirable to use only MDS assessments to capture in-SNF use of such services. A revision in the MDS assessment instrument that reports whether a patient received a service in a SNF would be required to implement this option efficiently.

Pros:

- It accounts for 2 important high NTA cost subgroups (IV therapy and RT)
- It has a reasonable amount of variance explanation of NTA costs (21%)
- It is based solely on SNF data
- It encourages delivery of certain services (e.g., RT) that might otherwise be underprovided because of high costs

Cons:

- Efficient administration of the SIM model is dependent on revisions to the MDS assessment instrument
- It is not strongly based on clinical/diagnostic information

New Profile NTA Model (NP-NTA)

Starting from scratch, the NP-NTA model develops a patient classification model for NTA costs that is most clinically meaningful, using variables that minimize gaming opportunities by SNF providers (See Chapter V). The model starts with the stratification of SNF patients into acute, chronic, and rehabilitation subgroups, characterizing the major types of Medicare SNF patients. Other variables are then incorporated to finalize the specific models for each of the NTA component cost categories.

The NP-NTA model includes 22, 28, and 30 patient adjustment factors for the acute, chronic, and rehabilitation subgroups, respectively. These factors are based on condition and service use variables from SNF claims, prior hospital claims, and MDS assessments. All variables are selected to have appropriate incentives for quality of care and to have limited game-ability.

Pros:

- It is designed specifically to be clinically meaningful
- It has the highest variance explanation of NTA costs (25%)
- It provides appropriate clinical incentives and encourages delivery of certain services (e.g., RT) that might otherwise be underprovided because of high costs

Cons:

- Like SIM, it is dependent on revisions to MDS
- Implementation/administration requirements are more complex than for the other approaches
- It requires data from prior hospital stays
- It may set up a communications barrier between hospital and SNF

Statistical Comparison of the three NTA Models

The NTA models are compared according to the multiple criteria noted above. We examine statistical indicators derived from a common sample using common methods, which we call “the base case,” to provide consistent comparisons. We compare the models on a person and on a facility basis. We also compare results for relative weights based on both NTA costs and charges. Because we disaggregated some NTA services (e.g., RT) to the department level and included a wide range of cost-to-charge ratios for them, we were uncertain about the validity of some of the cost-to-charge ratios used to convert claims-based charges to costs. While it does not provide validation for the cost-based models, examining models with charge-based relative weights provides a sense of the potential influence of extreme cost-to-charge ratios on model performance.

Comparison of Person-Level NTA Models Based on Costs

Table VII.1 presents statistics on our predictive models of person-level NTA costs. The upper panel presents findings from our test sample; the lower panel presents comparable information from the validation sample. Because the results are very similar, regardless of the sample used, we focus discussion on the validation sample results. The first comparison indicates the amount of variance in NTA costs explained by the models. RUG-58, the first of our three NTA models, explains about 10% of the variation in NTA costs, an amount that is considerably better than the 7% explained by the recalibrated RUG-III system currently in effect. The second model, the SIM and RUG-58G combination, accounts for 21% of the variation, twice that of RUG-58 alone. Since it is an extension of RUG-58, the SIM+RUG58G builds on the

Table VII.1: Comparison of Person-Level NTA Costs Models: RUG-III, RUG-58, SIM+RUG-58G[a], and NP-NTA

Indicators	RUG-III	RUG-58	SIM+RUG-58G[a]	NP-NTA
a. Testing sample				
R-squared	0.064	0.095	0.212	0.249
SD relative weights	0.354	0.429	0.641	0.680
Sensitivity	0.314	0.376	0.454	0.461
b. Validation sample				
R-squared	0.068	0.103	0.214	0.243
SD relative weights	0.349	0.426	0.626	0.670
Sensitivity	0.313	0.375	0.442	0.460

a. RUG-58G contains 8 categories indicating the first digit of the RUG-58 score.

effects of that model to explain NTA costs. Our third model, NP-NTA, has the highest R-squared, explaining almost 25% of the variation in NTA costs.

The second person-level statistic we examine is the standard deviation of the relative weights (SD-RW), which indicates the range of payment weights covered by the models. The SD-RW increases with the complexity of the models. For example, while RUG-58 has a SD-RW of 0.426, the SIM+RUG58G has a SD-RW of 0.626, roughly 50% higher. The SD-RW of the NP-NTA (0.670) is slightly higher even than that of the SIM+RUG58G, indicating that the latter model has a slightly greater spread in payment values.

The third person-level statistic we examine is the sensitivity statistic, which indicates the percentage of the highest cost cases correctly predicted by the model to be in the top 10% of predicted values. RUG-58 correctly predicts 37% of the highest (10%) cost cases, compared with 44% by SIM+RUG58G. The NP-NTA model scores the highest of the tested models, correctly predicting 46% of the highest cost cases.

Comparison of Person-Level NTA Models Based on Charges

Table VII.2 presents statistics that are analogous to those in Table VII.1, but for charges rather than costs. Some of the basic patterns we observed in Table VII.1 also appear in Table VII.2. For example, the R-squared increases monotonically from RUG-III to NP-NTA. The SD-RWs and the sensitivity indicators also increase with the complexity of the models, from RUG-III to NP-NTA.

Table VII.2: Comparison of Person-Level NTA Charge Models: RUG-III, RUG-58, SIM+RUG-58G[a], and NP-NTA

Indicators	RUG-III	RUG-58	SIM+RUG58G[a]	NP-NTA
a. Testing sample				
R-squared	0.059	0.097	0.283	0.362
SD relative weights	0.468	0.598	1.02	1.13
Sensitivity	0.300	0.321	0.516	0.549
b. Validation sample				
R-squared	0.063	0.104	0.259	0.342
SD relative weights	0.462	0.594	0.993	1.14
Sensitivity	0.294	0.320	0.489	0.491

a. RUG-58G contains 8 categories indicating the first digit of the RUG-58 score.

We also found notable differences between the models estimated with cost-based weights and those with charge-based weights. While the differences in R-squared between the cost-based and the charge-based models are marginal for the RUG-III and RUG-58 models, they are considerable for the other two models. For example, the R-squared of the charge-based NP-NTA model (0.34) was 10 percentage points higher than the R-squared of the cost-based NP-NTA model (0.24). In general, the SD-RWs are larger for the charge-based models than for their cost-based counterparts. One explanation for this difference is that dispersion charges also reflect variations in SNF charging, whereas this variability was appropriately constrained, in the case of costs, by the cost-to-charge ratios applied to charges. Another possible explanation is that the charge models fit better because the cost-to-charge ratios are assumed fixed across patients in the same facility, thereby masking true patient variation in resource use within facilities. The facility-level analyses below suggest that the former explanation is the primary reason for the difference.

In the cases of RUG-III and RUG-58, the sensitivity indicators were slightly higher for the cost-based models. For the SIM+RUG58G and NP-NTA models, in contrast, the sensitivity indicators were slightly higher for the charge-based models.

Comparison of Facility-Level Models Based on Cost Weights

Data from the validation sample for the person-level models were used to calculate facility-level models. Table VII.3 presents the evaluation statistics for facility-level models that were derived from cost-based relative weights (as in Table VII.1). The first panel presents the evaluation statistics for just the case-mix index. While RUG-58 explains only 14% of the variation in facility-level NTA costs, SIM+RUG58G explains 35%, a more than two-fold difference. The NP-NTA model has a slightly higher R-squared (36%) than the SIM+RUG58G model. In comparison, RUG-III only explains 5% of the variation in facility-level NTA costs. It is important to note that the RUG-III statistics shown below predicate what the classification *could do* if the categories were fit to costs, in contrast to the current system's use of nursing

Table VII.3: Comparison of Facility-Level NTA Costs Models: RUG-III, RUG-58, SIM+RUG58G[a], and NP-NTA

Indicators	RUG-III	RUG-58	SIM+RUG58G[a]	NP-NTA
a. CMI only				
R-squared	0.047	0.136	0.347	0.362
CMI coefficient	0.899	1.13	1.29	1.15
SE on CMI coefficient	0.125	0.092	0.051	0.041
p-value	0.419	0.155	0.000	0.010
b. CMI payment model[b]				
R-squared	0.066	0.153	0.353	0.400
CMI coefficient	0.958	1.15	1.29	1.15
SE on CMI coefficient	0.124	0.091	0.051	0.040
p-value	0.737	0.095	0.000	0.010
c. CMI full model[c]				
R-squared	0.263	0.295	0.396	0.421
CMI coefficient	0.627	0.756	1.040	0.995
SE on CMI coefficient	0.105	0.082	0.064	0.047
p-value	0.000	0.003	0.486	0.917

a. RUG-58G contains 8 categories indicating the first digit of the RUG-58 score.
b. The payment model includes urban/rural status and the area wage index, in addition to the CMI.
c. The full model includes urban/rural status, area wage index, indicators for number of SNF beds (0-9 beds, 10-19 beds, 20-49 beds, 50-99 beds, with 100+ beds as the excluded category), hospital based dummy, ownership type (voluntary/nonprofit dummy, for profit other dummy, and government dummy, with for profit chain as the excluded category) and percent Medicare days, in addition to the CMI.

weights for NTA costs.

In the first panel of Table VII.3, the CMI coefficient for the RUG-58 model (1.13) and NP-NTA model (1.15) reveal only small increases from the desired 1.0 level. The CMI coefficient for the SIM+RUG58G model (1.29) is much higher. For the NP-NTA model, the coefficient is 1.15. We performed tests of statistical significance to determine whether the CMI coefficient is statistically different from 1.0. For RUG-58, the CMI (at 1.13) is not significantly different from 1.0 (p=.155). In contrast, the CMI for SIM+RUG58G (at 1.29) is significantly different from 1.0, as is the CMI coefficient of NP-NTA (at 1.15).

The second panel in Table VII.3 presents statistics for the preferred models when conventional Medicare payment adjusters—urban/rural status and wage index—are included with the CMI. Addition of the two variables does not notably affect the evaluation statistics shown in the first panel.

The third panel presents findings when a number of other facility characteristics are included in the models. These include: (a) hospital-based status, (b) percent Medicare, (c) control of ownership, (d) number of SNF beds, and (e) percent of total days covered by Medicare. For the RUG-58 model, controlling for the other factors strongly affects the evaluation statistics. Variance explanation (R-squared), for example, increases to almost 30%, indicating that the facility characteristics added to the model explain a considerable amount of the variation in facility-level NTA costs. The coefficient on the CMI drops notably below 1.0. With the additional variables, the coefficient on the CMI of the SIM+RUG58G model drops to 1.04, not significantly different from 1.0 (p-value=.486). Similarly, the CMI coefficient on the NP-NTA model drops to 0.995, also not statistically different from 1.0. For the RUG-III model, the variance explanation rises to 26%, but the CMI coefficient is only 0.627, a statistically significant indicator of decompression.

Although not presented in Table VII.3, we conducted a similar evaluation of the CART version of the SIM+RUG58G cost model. Findings mirror those of the regression models presented here. In general, the amount of variance explained by the CART version is slightly smaller than that of the regression version, while the CMIs are generally higher.

Comparisons of Facility-Level NTA Models Based on Charge Weights

Table VII.4 presents analogous statistics to those in Table VII.3, except that the models were estimated with charge-based relative weights (as in Table VII.2). We highlight the notable differences between the two tables.

Table VII.4: Comparison of Facility-Level NTA Charges Models: RUG-III, RUG-58, SIM+RUG-58G[a], and NP-NTA

Indicators	RUG-III	RUG-58	SIM+RUG-58G[a]	NP-NTA
a. CMI only				
R-squared	0.048	0.144	0.355	0.312
CMI coefficient	0.641	0.794	0.866	0.554
SE on CMI coefficient	0.090	0.062	0.033	0.022
p-value	0.000	0.001	0.000	0.010
b. CMI payment model[b]				
R-squared	0.066	0.161	0.359	0.349
CMI coefficient	0.682	0.806	0.859	0.553
SE on CMI coefficient	0.089	0.061	0.033	0.022
p-value	0.000	0.002	0.000	0.010
c. CMI full model[c]				
R-squared	0.256	0.291	0.396	0.378
CMI coefficient	0.369	0.506	0.706	0.457
SE on CMI coefficient	0.078	0.058	0.043	0.025
p-value	0.000	0.000	0.000	0.010

a. RUG-58G contains 8 categories indicating the first digit of the RUG-58 score.
b. The payment model includes urban/rural status and the area wage index, in addition to the CMI.
c. The full model includes urban/rural status, area wage index, indicators for number of SNF beds (0-9 beds, 10-19 beds, 20-49 beds, 50-99 beds, with 100+ beds as the excluded category), hospital based dummy, ownership type (voluntary/nonprofit dummy, for profit other dummy, and government dummy, with for profit chain as the excluded category) and percent Medicare days, in addition to the CMI.

The amount of variance in NTA costs explained by each model is about the same, although the NP-NTA charge models tend to have noticeably higher R-squared than their cost-based models. The most notable difference between the models in Table VII.3 and VII.4 is that the CMI coefficients in the charge-based models are uniformly less than 1.0, whereas most of the CMIs in the cost-based models are greater than 1.0 and only sometimes significant. In the case of the CMI payment model, for example, the CMI coefficient for the SIM+RUG58G model is 1.29 for the cost-based model and 0.85 for the charge-based models. If relative weights were based on costs, these findings suggest that high CMI providers would receive payments lower than costs, and low CMI providers would receive payments higher than costs. If relative weights were based on charges, high CMI providers would receive payments higher than costs and low CMI providers would receive payments lower than costs. These latter findings seem to be a reflection of the different charging practices across SNFs, as indicated by the higher standard deviation of the relative weights of the charge-based models.

Summary of Statistical Comparisons

In sum, the amount of NTA cost variance explained by the RUG-58 model (10%) is higher than that of RUG-III (7%). While RUG-58 represents an improvement from the current system, the SIM and NP-NTA models are both better, raising the R-squared for NTA costs to 21% and 25%, respectively. The latter two models have levels of variance explanation that are in the "ball park" with the R-squared derived for case-mix classification systems for Medicare payment for long-term care hospitals and psychiatric hospitals, although they are considerably lower than the R-squared derived for inpatient rehabilitation hospitals. In addition, RUG-58 correctly predicts 38% of the highest cost (10%) cases, whereas SIM and NP-NTA correctly predict about 45% of those cases.

The facility regressions provide mixed information on whether compression is a problem. When considering the payment models (including CMI, wage index, and urban/rural status), the CMI for RUG-58 (1.15) is not significantly different from 1.0; however, the coefficients are statistically significant for SIM (1.29) and NP-NTA (1.15). These results suggest some compression problems with both of the more complex classification models, although in neither case is the CMI coefficient dramatically greater than 1.0. When other facility characteristics are included in the full model, the CMI coefficient for RUG-58 drops to a statistically significant

0.76, compared with 1.04 for SIM, which is insignificantly different from 1.0. The CMI coefficient for NP-NTA drops to 0.995, also not significant. Under the payment models, the results suggest that all three models are likely to underpay SNFs with more difficult case-mixes for NTA services, relative to their costs. Under the full model, the results suggest that RUG-58 may overpay such facilities relative to their costs, whereas the other two models pay facilities nearly in line with their costs. Notwithstanding, all three models are would pay closer in proportion to costs than the current RUG-III system.

Comparisons of NTA Models Based on Other Policy Criteria

Other criteria that we examined were: (a) clinical meaningfulness, (b) ease of implementation and operation, (c) ease of use by providers, (d) and appropriate clinical and financial incentives. We compare the three options relative to these criteria in turn:

Clinical Meaningfulness

The NP-NTA model is the most clinically meaningful of the three since it was designed to characterize the broad categories of the Medicare SNF population, as well as the subgroups within these categories. SIM is slightly more clinically meaningful than the RUG-58 because it specifically identifies two subgroups—users of RT and IV medications—that have notably high NTA costs.

Ease of Implementation and Administration

RUG-58 is the easiest model to implement because it entails a straightforward expansion of RUG-III groups, based on the interaction of groups used in the current system. SIM is slightly more complicated to implement. With modifications to the MDS assessments, SIM can be implemented using a few additional MDS variables. NP-NTA also requires modification to the MDS, plus a mechanism to convey information from prior hospital stays to supplement SNF stay-specific information. While some information is currently transferred from discharging hospitals to SNFs, the amount and types of data are variable. Hence, a systematic process with very specific minimum data items would have to be developed.

Ease of Use by Providers

RUG-58 places no additional burden on providers, since it uses only the data already required by the current SNF PPS. Additionally, SNF providers are already familiar with this model. Although SIM is slightly more complicated than RUG-58, it uses only SNF generated data. The proposed changes to ensure certain services are provided during SNF stays would be viewed as a common sense addition to the MDS. In comparison, the NP-NTA model is more complicated than SIM, and requires information from prior hospital stays to fully characterize the payment groups. If hospital stay information can efficiently be transmitted to SNFs, the use of NP-NTA would also be straightforward and particularly useful to smaller SNFs, which may not currently have systematic procedures for retrieving data from prior hospital stays.

Appropriate Clinical and Financial Incentives

All three models include service use indicators. This is a potential concern; when payment is related to service use (rather than the need for services), opportunities arise for inappropriate clinical and financial incentives. Although we selected and defined variables that were relatively ‘less risky’ for gaming in all three models, the fewer the number of services included, the less risk for inappropriate incentives. At the same time, it should be noted that inclusion of some service use variables in payment models would provide incentives to provide certain types and levels of care (e.g., RT care by licensed therapists) that might be more efficiently provided in SNFs than in other Medicare provider settings.

C. Rehabilitation Therapy Specific Options

New Profile Rehabilitation Therapy Model (NP-Therapy)

We developed one rehabilitation therapy-specific option based on patient need for therapy services: NP-Therapy (see Chapter V). This “need-based” option addresses the concern that rehabilitation therapy services are paid on a fee schedule in the current SNF PPS. Under the current system, providers have considerable latitude to determine how much therapy is provided, and may respond to financial incentives as well as to patients’ needs. Adoption of the NP-Therapy model by itself implies that the rehabilitation therapy component in the current SNF PPS would be replaced, while the rest of the payment system would continue to be used as is. Only the rehabilitation therapy payments in the current SNF PPS would be re-distributed according to this model.

The NP-Therapy model starts with the stratification of Medicare SNF patients into three basic subgroups; acute, chronic, and rehabilitation. Other variables interact with these groups in the final model. The NP-Therapy model includes multiple patient condition variables from SNF, prior hospital claims and MDS. The vast majority are MDS patients' functional status variables.

Pros:

- It is need based
- It is clinically meaningful
- It provides reasonable variance explanation (19%)
- It could reduce over-provision of therapy services

Cons:

- There is a risk of providers under-providing services, relative to the fee schedule
- If there is under-provision, this model could exacerbate it, given the paucity of information about the right amount of therapy services to provide any given patient
- As with any classification change, this one implies new implementation and administration requirements
- It requires transmission of data from prior hospital stays

Statistical Comparison of "Need Based" and RUG-III Rehab Therapy Options

The major comparison among options for dealing with rehabilitation therapy is between the NP-Therapy models, which were developed separately for PT/OT and ST, and the corresponding RUG-III rehabilitation therapy components. We applied the same indicators for comparing the rehabilitation therapy options that we used to examine NTA options.

Comparison of Person-Level Rehabilitation Cost Models

Table VII.5 presents statistics on our predictive models of person-level rehabilitation therapy costs. The upper panel presents findings from our test sample; the lower one presents comparable information from the validation sample. Differences in evaluation statistics between the two are small, so we focus discussion on the latter. The first column presents statistics for the RUG-III model for physical therapy (PT)/occupational therapy (OT); the second column presents comparable statistics for the NP-Therapy (PT/OT) model. The third and fourth columns present information for comparisons between the RUG-III and NP-Therapy speech therapy (ST) models. We chose to keep these two types of rehabilitation therapy separate, at least at present, because the patient characteristics predicting the two types are quite different. These components can be combined into a single model with further development work.

Table VII.5: Comparison of Person-Level Rehabilitation Therapy Costs Models: RUG-III and NP-Therapy

Indicators	PT/OT		ST	
	RUG-III	NP	RUG-III	NP
a. Testing Sample				
R-squared	0.386	0.192	0.108	0.191
SD relative weights	0.492	0.345	0.987	1.315
Sensitivity	0.336	0.277	0.485	0.424
b. Validation Sample				
R-squared	0.387	0.193	0.098	0.176
SD relative weights	0.490	0.341	0.990	1.32
Sensitivity	0.342	0.293	0.469	0.426

For the PT/OT models, the amount of variance in rehabilitation therapy costs (R-squared) explained by the RUG-III system (39%) is about twice as high as those for the NP-Therapy (PT/OT) mode (19%). This result is not surprising, however, since the RUG-III rehabilitation therapy component was designed to categorize cells according to the number of minutes of therapy provided (or expected to be provided). Thus, the high R-squared for this model is due in large part to the relationship between minutes of therapy provided according to the payment schedule and costs of providing those services. RUG-III has a SD-RW of 0.49, while NP-Therapy (PT/OT) has a SD-RW of only 0.35, indicating that RUG-III has a wider range of payment weights for PT/ OT therapy services. Although the sensitivity indicator is slightly higher for RUG-III (34% vs. 29%), both models correctly predict about 30% of the highest cost cases.

For ST services, the amount of variance explained by the NP-Therapy (ST) model is almost twice that of the RUG-III speech therapy model (18% vs. 10%). These statistics are from the validation sample, which provides a better representation of expected results than the test sample that was used to develop the models. NP-Therapy (ST) model has a SD-RW of 1.32, which is considerably wider than that of the RUG-III model (0.99). The large SD-RW may be due, in part, to the large number of stays with zero ST use. In contrast to their differences on the SD-RW, the two models have very comparable values on the sensitivity indicator (47% for RUG-III and 43% for NP-Therapy (PT/OT)).

Comparison of Facility-Level Rehab Cost Models

Data from the validation sample for the person-level models were used to calculate facility-level models. Table VII.6 presents the evaluation statistics, similar to those we presented above for the NTA models. We focus on the payment model and the full model.

Table VII.6: Comparison of Facility-Level Rehabilitation Therapy Costs Models: RUG-III and NP-Therapy

Indicators	PT/OT		ST	
	RUG-III	NP	RUG-III	NP
a. CMI only				
R-square	0.313	0.139	0.142	0.195
CMI coefficient	0.889	1.20	1.03	1.02
SE on CMI coefficient	0.070	0.110	0.074	0.053
p-value	0.114	0.073	0.715	0.675
b. CMI payment model[a]				
R-square	0.329	0.149	0.147	0.203
CMI coefficient	0.910	1.20	1.01	1.01
SE on CMI coefficient	0.071	0.107	0.075	0.053
p-value	0.207	0.061	0.946	0.793
c. CMI full model[b]				
R-square	0.373	0.171	0.168	0.210
CMI coefficient	0.914	1.200	0.945	0.970
SE on CMI coefficient	0.068	0.120	0.081	0.058
p-value	0.209	0.103	0.500	0.550

a. The payment model includes urban/rural status and the area wage index, in addition to the CMI.
b. The full model includes urban/rural status, area wage index, indicators for number of SNF beds (0-9 beds, 10-19 beds, 20-49 beds, 50-99 beds, with 100+ beds as the excluded category), hospital based dummy, ownership type (voluntary/nonprofit dummy, for profit other dummy, and government dummy, with for profit chain as the excluded category) and percent Medicare days, in addition to the CMI.

For physical therapy and occupational therapy services, the RUG-III payment model explains 33% of the variation in rehabilitation therapy costs at the facility level, while NP-Therapy (PT/OT) explains 15% of the variation in costs. The higher R-squared of the RUG-III facility-level model, relative to the NP-Therapy model, is again due largely to the circularity between RUG-III payment categories and costs. The CMI coefficient of RUG-III is 0.91 which is not significantly different from 1.0. The NP-Therapy (PT/OT) has a CMI coefficient of 1.2; although not significantly different from 1.0 (p=.061) at the .05 level, it is significantly different at the .10 level. In the full models, R-squared values for both RUG-III and NP-Therapy models do not change appreciably relative to those of the payment models. The values of the CMI coefficients are also unchanged.

For the payment models for speech therapy, the NP-Therapy (ST) model explains a higher proportion of the costs than does the RUG-III model (20% vs. 15%). Both models have a CMI coefficient that is effectively 1.0. Results for the full models are also quite similar for the two models. NP-Therapy (ST) has a slightly higher R-squared (21%) than does the RUG-III model (17%), and both models have CMI coefficients that are close to 1.0 and not significantly different from 1.0.

In sum, the RUG-III model explains a higher proportion of the variance in rehabilitation costs for physical therapy and occupational therapy, at both the person and facility levels, but a large part of its R-squared value is due to the relationship between the RUG-III payment categories and rehabilitation costs. RUG-III also has a CMI coefficient slightly below 1.0 and insignificant. NP-Therapy (PT/OT) has a CMI coefficient slightly above 1.0 but also insignificant. For speech therapy, NP-Therapy (ST) explains a higher proportion of the variance in costs than does the RUG-III at both the person- and the facility-level. For the payment model, the CMI coefficients for both models are effectively stay at 1.0. In general, neither model is superior to the other.

Comparison of NP-Therapy Models and RUG-III Based on other Policy Criteria

Clinical Meaningfulness

NP-Therapy was designed to be clinically meaningful to practitioners, and the first level of the classification model is based on the three major categories of Medicare SNF patients. Within each of the broad categories, subgroups are defined primarily by functional status indicators, which suggest need for therapy services based on clinical criteria. The RUG-III therapy component is composed of bands of therapy minutes provided (or predicted to be provided), modified by therapist type. This component does not include indicators of patients' clinical or functional status.

Ease of Implementation and Administration

NP-Therapy can be implemented with some changes to the information systems of CMS and providers, since most of the variables in the model are derived from the MDS. It is necessary, however, to implement a process for transmitting data from prior hospital stays to complete the classification of patients.

Ease of Use by Providers

Providers have been adjusting to use the RUG-III based classification system since 1998. Naturally, they would probably think it desirable to continue its use. New adjustments would be needed if NP-Therapy were implemented, since each SNF provider would have to determine how the new payment rates relate to the costs of the therapy services it actually provides.

Appropriate Clinical and Financial Incentives

The current rehabilitation therapy component of the SNF PPS gives providers considerable latitude over the amount and type of therapy services provided to patients. While some observers feel that financial incentives are too strong under the fee schedule payment system, some providers think the RUG-III rehabilitation component fosters a "case management" approach by allowing more flexibility in adapting care to changing needs. Although a "need-based" rehabilitation therapy model is more conventional, it may create incentives to *under-provide,* since there is no established 'right amount' of therapy to provide a given patient. Some observers argue that the fee-schedule provides more assurance that needed therapy is actually provided.

D. DRG-Related Options

Unlike the NTA-specific and rehabilitation therapy-specific options—which were designed for SNF services—the DRG-based option was "taken off the shelf," with enhancements for functional and cognitive status, and examined for its applicability to SNF service costs.

Adoption of the DRG total cost model would imply implementation of a totally different approach for case-mix classification of Medicare SNF services. As described in Chapter VI, the DRG classification system also can be adopted to case-mix adjust payments for NTA services, rehabilitation therapy services, or both, with the remaining service components continuing to be paid under the RUG-III system.

Model

Conceptually, the DRG-based system begins to address some of the areas for potential refinement in the current SNF PPS. It improves the variance explanation for NTA costs, substitutes a need-based payment strategy for rehabilitation therapy services, and provides a possibility for updating payment weights in the fashion used for acute care hospitals under Medicare. Although we tested a DRG-based model using the prior hospital stay data to assign DRGs, it is also possible to assign SNF patients to DRGs based only on SNF data, once SNF coding issues are resolved (see Chapter III). In either case, once DRGs are assigned, weights are constructed based on SNF cost or charge data. Our model includes a small set of functional status indicators to augment the DRGs. Other patient characteristics information from the SNF stay could readily be incorporated to further enhance the DRG-based model.

Pros:

- The model relies only on DRGs from the prior hospital stays and functional status data from the MDS
- It can be used to pay on the basis of total costs of care
- It is not subject to gaming by SNF providers

Cons:

- Patient conditions that could change between hospital and SNF stays would not be reflected
- It requires data from prior hospital stays
- It only marginally captures medical complexity during SNF stays
- The hospital used for the prior stay could influence SNF payment

Statistical Comparison of DRG-Based and RUG-III Models

Comparison of Person-Level Cost Models

Table VII.7 presents statistics on the DRG-based person-level models for total service costs. (See Chapter VI for statistical evaluations of the NTA and rehabilitation therapy components.) We provide comparable statistics for the RUG-III model in two ways.

First we used RUG-III categories to explain total costs of care. While the RUG-III system was not originally designed for this purpose, we illustrate what the RUG-III classification system could do if the RUG weights were redefined to predict SNF costs (rather than predicting staff nursing time and pricing actual therapy minutes). It is analogous to the other comparisons to RUG-III above for the different cost components.

Second, we used the current RUG-III relative Medicare SNF payments to predict total SNF costs. This is meant to illustrate in a comparable manner how well the current RUG-III payments relate to total SNF costs. We implement this analysis in two steps. First, we compute the average total SNF payments by RUG group by regressing actual per day payments on the 44 groups, from which we compute RUG payment weights. We then regress total costs per day on the payment weights at the stay level. We also use the RUG-III payment-based relative weights to construct CMIs for the facility-level analyses that follow. Because payments are not directly available for components of services, we limited the latter comparison to one of total Medicare service payments.

Table VII.7: Comparison of Person-Level SNF Total Costs Models, RUG-III and DRG + Functioning

Indicators	Explanatory Variables		
	RUG-III Categories	RUG-III Payment Weights	DRG + Functioning Variables
a. Testing sample			
R-square	0.044	0.019	0.128
SD relative weights	0.115	0.134	0.190
Sensitivity	0.221	0.103	0.248
b. Validation sample			
R-square	0.050	0.021	0.109
SD relative weights	0.116	0.134	0.189
Sensitivity	0.226	0.098	0.231

As with other comparisons in this chapter, the results for the validation sample are similar to those for the testing sample, so we focus discussion on the lower panel of Table VII.7. The results in the first two columns compare how RUG-III predicts costs with how the RUG-III payment weights currently predict costs. The R-squared is 5.0% using the RUG-III categories versus 2.1% using the RUG payment weights. Thus, the 44 RUG groups could predict costs better than RUG payments currently do. The 44 RUG groups could predict 23 % of the most costly 10% of cases, whereas current RUG-payments only predict 9.8% of the most costly cases. Interestingly, RUG-III payment would vary more if based on relative costs. The bottom line is that current payments vary by RUG group in ways that bear little relation to variation in patient costs. Even if the RUG-III weights were replaced with cost-based weights, the R-squared would only be 5.0%.

Comparing the DRG + functioning models in the third column to the RUG-III models in the first column, we see that the DRG-based model explains 11% of variation in costs and correctly predicts 23% of the highest (10%) cost cases. RUG-III explains only 5% of the variance in total costs, but correctly predicts 23% of the highest cost cases. The standard deviation of the relative weights of the DRG model is 0.189, relative to 0.116 for RUG-III, indicating that the DRGs would have a wider range of payment weights for total costs. This is due to the ability of DRGs to better capture routine costs compared with RUG-III.

Comparison of Facility-Level Cost Models

Table VII.8 presents evaluation statistics for facility regressions where total facility costs are regressed on three different CMI measures: the RUG-III cost based CMI, RUG-III current payment-based CMI, and DRGs + functioning cost-based CMI. As for other analyses above, CMI only models, payment models, and fully-specified models are estimated.

Table VII.8: Comparison of Facility-Level Total Costs Models: RUG-III and DRG + Functioning

	CMI Measure		
	RUG-III Cost-Based CMI	**RUG-III Current Payment-Based CMI**	**DRGs + Functioning Cost-Based CMI**
a. CMI only			
R-square	0.136	0.051	0.248
CMI coefficient	2.40	1.19	1.998
SE on CMI coefficient	0.232	0.126	0.116
p-value	<0.005	0.137	<0.005
b. CMI payment model			
R-square	0.192	0.104	0.289
CMI coefficient	2.46	1.22	1.955
SE on CMI coefficient	0.233	0.127	0.115
p-value	<0.005	0.085	<0.005
c. CMI full model			
R-square	0.592	0.594	0.577
CMI coefficient	1.06	0.883	0.467
SE on CMI coefficient	0.155	0.108	0.096
p-value	0.677	0.281	<0.005

The RUG-III cost-based CMI models have higher R-squared than the payment-based CMI models in the CMI-only and payment regressions (13.6% vs. 5.1% and 19.2% vs. 10.4%, respectively). In the fully-specified models, however, the R-squared are nearly identical in the 2 models (about 59%). This suggests that much of the additional explanatory power exhibited by cost-based CMIs is related to facility characteristics. In addition, the high R-squared in the facility regressions suggests that facility characteristics, such as ownership, hospital-based status, and bed size, explain a large share of the variation in total SNF costs.

We find evidence of extreme CMI compression in the payment model using the RUG-III cost CMI, but only moderate compression using the payment-based CMI. This suggests that current payments are more proportionate to costs than are RUG-III cost-based relative weights. In the fully-specified models, however, we see that the RUG-III cost-based models exhibit only minor (and statistically insignificant) compression; the payment-based CMI looks as if it exhibits CMI decompression (CMI coefficient = 0.888). Due to the large standard error, we cannot conclude it is significantly different from 1.0; but the point estimate suggests that under the current RUG-III payment regime, facilities may in fact have incentives to select patients with high expected costs. The rehabilitation RUG categories tend to be the highest paying, and some observers consider them to be the most profitable categories for SNFs. This would be consistent with the decompression point estimate.

Comparing the DRG + functioning cost-based CMI results with the RUG-III cost-based CMI results, we find that in the CMI only and payment regressions, the DRG-based CMIs explain a larger share of total facility costs than the RUG-III CMI (24.8% vs. 13.6% and 28.9% vs. 19.2%, respectively). In the fully-specified models, however, the R-squared is slightly lower using DRGs compared with RUGs (57.7% vs. 59.2%). This suggests the RUGs are picking up more unique cost variation not otherwise explained by facility characteristics. The DRG CMI, like the RUG-III CMI, exhibits extreme compression in the payment regression.

In the fully-specified models, the DRG CMI again exhibits substantial and statistically significant decompression, with a CMI coefficient of 0.467. This suggests that using the DRG-based model without further modifications would create a larger incentive for SNFs to select patients in the higher-paying case-mix groups than is currently exhibited for the RUG-III system or for a RUG-III system with cost-based relative weights.

Comparison of DRG-Based Models and RUG-III Models Based on Other Policy Criteria

Clinical Meaningfulness

The DRG-based model that we developed uses the DRGs from prior hospital stays, along with functional and cognitive status measured during the SNF stay. Because the SNF stay is conceptually an extension of the hospital stay, the DRG of the prior hospital stay would differentiate, to some extent, the high cost of SNF services. In addition, the DRG model includes measures of functional status during the SNF stay. A missing piece in this model is the medical complexity of patients during the SNF stay, which may include, for example, use of IV medication or RT. Because of its particular use of extensive services, clinically complex, and special care categories, RUG-III more directly captures some of this medical complexity.

Ease of Implementation and Administration

Because DRGs are in current use for acute and long-term care hospitals, most of the process machinery to implement a DRG-based model is readily available. Relative to RUG-III, some additional effort is required to move to a new system, but the effort should be modest. Nevertheless, some systematic means for transmitting DRG information from prior hospital stays to SNFs would have to be implemented.

Ease of Use by Providers

SNF providers are generally unfamiliar with DRGs, but are quite familiar with the functional status measures from the MDS. Adjusting to DRGs, based on prior hospital stays, will be a burden for providers in the short run. In the long run, however, it seems reasonable to transition to using diagnosis information from SNF stays as input to the DRG-based classification system.

Appropriate Clinical and Financial Incentives

The DRG-based model is based on DRGs of prior hospital stays and current functional status variables. It does not, therefore, contain incentives to provide particular services for financial, rather than clinical, reasons. Although some observers claim that DRGs based on prior hospital stays present opportunities for acute care hospitals to influence SNF payments, others argue that it is not cost-effective for such hospitals to modify DRG coding simply in order to influence SNF payments.

E. Discussion

Since there are pros and cons for each of the models, choice of a specific model will depend on how much weight is placed on particular criteria used for comparing the models.

The models we have presented in this report reflect particular versions of research options for addressing specific SNF service components. The presented models can certainly be refined, based on further input. For example, specific variables could be deleted from a given model if they are deemed to provide inappropriate incentives. Also, although we presented multiple models that span a continuum of choices according to varying degrees of statistical, clinical and administrative complexity, hybrids of those models can be easily developed.

VIII. SNF PPS Outlier Payment Policies

A. Introduction

Medicare outlier payments are additional compensation to providers paid under prospective payment systems (PPSs) for cases whose treatment costs are extraordinarily high, as determined by specific threshold criteria, relative to their regular PPS payments. Outlier payments are budget-neutral policies, meaning that they are self-financed through reductions made to the PPS's base payment rate.

The SNF PPS currently does not include an outlier policy. However as this project's TAP members noted (SNF TAP 2004), some stakeholders and analysts suggest that an outlier policy should be a basic component of any PPS, and others suggest that an outlier policy may help address the SNF PPS case-mix classification system's specific limitations in matching NTA payments with NTA resource use—a limitation that became apparent only after the PPS had been introduced.

Outlier payment policies are not sufficient, however, for fully realigning mismatches between payments and costs that occur due to the structure of the classification system, and they cannot turn a patient stay or patient day from a financial loss into either a financial draw or a financial gain. Instead, outlier payments defray only *some* of the losses incurred by providers for a given case, and are intended to reduce a PPS's incentive for providers to undertreat such cases. Provider losses before outlier payments are triggered are necessary to retain the efficiency incentives of a PPS (Keeler, et al. 1988). Outlier policies are structured to compensate for the “upper” portion of a case’s loss, and require that the provider bear the “lower” portion of a case’s loss—that is, costs incurred above the PPS payment amount up to a specific threshold. This structure ensures that extreme costs are compensated for, but also results in many cases being ineligible for outlier payments, even though they incur some financial loss.

Outlier payment policies in each of Medicare’s other PPSs are based on an assessment of a case’s total costs in relation to its total PPS payment rate. This is the conventional approach because any type of patient care cost (e.g., prescription drugs, therapies, other ancillaries, nurse staffing) can contribute to the occurrence of extraordinary costs. (Patient care costs are identified by reducing patient care charges on a Medicare claim to costs using a provider’s ratio of costs to charges as determined from its Medicare cost report). Most TAP members and policy advisors to this project preferred, in principle, a total cost policy. But since the accuracy of the current methodology for allocating NTA costs has been a specific concern in the SNF PPS (Maxwell, et al. 2003), they expressed interest in an NTA-specific outlier policy as well. Further, some suggest that a narrower outlier policy may be less easily manipulated than a broader outlier policy, since providers have more latitude in manipulating multiple (rather than just one) service components. This occurred recently under the inpatient hospital PPS (Institute for Health and Socio-Economic Policy 2002) and corrective actions were issued in the hospital PPS proposed and final rules (CMS 2003). In this project, we developed and simulated both total cost outlier policies and NTA cost outlier policies. Under our NTA policy development and simulations, a case may not receive NTA outlier payments if its total PPS payments exceeded its total costs.

The rest of this chapter comprises seven sections. Section B explains in more detail the basic components of an outlier policy. Section C describes the data sources used and variables created. Section D describes the outlier policy approaches and parameters selected for developing and simulating specific outlier policies. Section E presents findings of the total cost policy simulations. Section F presents findings of the NTA cost policy simulations and broadly compares the Sections E and F findings of the two approaches. Section G makes additional comparisons between the two approaches by describing the "overlap" of the number of stays receiving outlier compensation under the two approaches, and by highlighting the two particular total and NTA policies that are most analogous to each other. Section H summarizes the chapter and concludes with tentative overall comments regarding outlier policies under the RUG-III SNF PPS.

Outlier policies are compared and discussed in terms of the incidence of outlier cases and outlier payments across patient and facility characteristics. The patient characteristics examined are the seven overall categories and forty-four final payment groups of the PPS's RUG-III case-mix classification system. The facility characteristics examined include free-standing or hospital-based status; rural or urban status; ownership and control status (non-profit, government, for-profit/chain-owned, or for-profit/ individually-owned); facility size (number of Medicare-certified beds); Medicare volume (percent of patient days that are Medicare-covered days); and the nine Census divisions.

The findings reveal that the incidence of both total cost and NTA cost outlier cases is highest in the RUG-III special care category relative to other RUG-III categories. Across all forty-four RUG-III classification groups, the incidence of outlier cases is highest in the rehabilitation low/B (RLB), special care/A (SSA), and extensive services/3 (SE3) groups. Most stays with outlier payments are in the rehabilitation category, though, reflecting the fact that most SNF stays are assigned to this RUG-III category. In terms of facility characteristics, outlier payments improve the average payment-to-cost ratios of hospital-based SNFs, government-owned SNFs, those with greater than a 25% Medicare volume share, and SNFs in West South Central states, relative to stays in facilities with other characteristics. Even among these particular facility groups, however, the outlier policies simulated do not alter the groups' average payment-to-cost ratios so much that they are close to 1.0.

B. Components of Medicare PPS Outlier Payment Policies

An outlier policy comprises three basic components, which finance and identify the cases that will receive outlier payments: (1) a targeted aggregate amount of funds made available for outlier payments; (2) a loss amount, which is used to determine an outlier eligibility threshold; and (3) and a loss-sharing ratio (i.e., a reinsurance rate).

Outlier Target

Outlier payments are financed by first targeting for each payment year a certain percentage of the total projected Medicare payments for a provider group (e.g., SNFs). Under some Medicare PPSs, this percentage is determined by statute. Up to this percentage of total spending can be used for outlier payments. However, outlier targets do not reflect designated dollars owed providers regardless of whether actual outlier payments fall short of the outlier target. Across Medicare PPSs, outlier targets range from 2% to 8% of aggregate Medicare provider group payments (Table VIII.1). The base payment rate (or conversion factor) in a PPS is reduced to finance the outlier target and still maintain budget neutrality.

Table VIII.1: Medicare Prospective Payment Systems: Outlier Policy Parameters, 2004

Provider Group's PPS	Unit of PPS Payment	Annual Outlier Target	Loss-Sharing Ratio	Loss Amount 2004
Inpatient (acute) hospital	Discharge	5%	80%	$31,000
Inpatient rehabilitation facility	Discharge	3%	80%	$11,211
Long-term hospital	Discharge	8%	80%	$19,590
Inpatient psychiatric facility	Day	2%	80% (1-8 days) 60% (9+ days)	$4,200 per stay
Hospital outpatient department	Bundle of services	3%	50%	2.6 x base rate
Home health agency	60-day period	5%	80%	1.13 x base rate

Loss Amount

Examining a case's total reported costs for services covered under a PPS rate identifies outlier cases. Some high cost, low prevalence services are excluded from PPS rates and are paid for separately.[24] As noted above, costs are identified by applying a provider's cost-to-charge ratio (CCR), derived from that provider's most recent Medicare cost report, to a case's total charges for covered services. In most of the Medicare PPSs, a provider's CCR that diverges very substantially from the CCRs of other similar facilities is replaced with an average CCR derived from those similar facilities (CMS 2003). A CCR may be very divergent because of, for

[24] Several high cost and low prevalence covered SNF services have been excluded from the SNF PPS daily rate and are paid for separately using Medicare's Part B fee schedules. These services include certain dialysis-related services and drugs, as well as services provided by physicians, midwives, psychologists, and nurse anesthetists (all excluded by the BBA). Also excluded are certain outpatient services when provided in a hospital (including associated medically indicated ambulance transport): cardiac catheterization, CT scans, MRIs, ambulatory surgery performed in operating rooms, emergency services, radiation therapy, angiography, and lymphatic and venous procedures (all excluded by the July 1998 Final Rule). Also excluded are: specified chemotherapy items and services, radioisotope services, customized prosthetic devices, and ambulance transportation for dialysis (all excluded by the BBRA).

example, data problems in cost reports, variations in cost allocation methods, or evidence of cost report manipulation (Institute for Health and Socio-Economic Policy 2002, CMS 2003). Analyses of the SNF cost reports under this project suggest that SNF cost report data issues may prompt the need for applying representative CCRs rather than actual facility CCRs in some cases.

A case qualifies for outlier payments when its costs for covered services exceed a specified dollar threshold. The threshold is determined by summing a case's PPS payments and a dollar amount, termed a loss amount. A *fixed*-loss amount reflects the level of unreimbursed costs that providers must bear for *any* case— before outlier payments can be received for costs incurred above the unreimbursed costs (i.e., above the fixed-loss amount). Fixed-loss amounts vary slightly by provider, because they are wage-adjusted to reflect labor costs in the area of the provider. As noted, loss amounts are necessary to retain the efficiency incentives of a prospective payment system (Keeler, et al. 1988). A fixed-loss amount, as opposed to a varying-loss amount, is intended to reduce a provider's differential incentives to accept cases in different case-mix classification groups. However, varying-loss amounts can be used by applying a specified multiplier to each case-mix classification group's payment rate. This multiplier approach is currently used in the hospital outpatient payment system, and was used in the hospital inpatient PPS until 1995.

Loss-Sharing Ratio

Providers are not reimbursed for 100 percent of reported costs above the threshold. Instead, outlier payments are intended to cover roughly the *marginal* costs incurred by providers (beyond the eligibility threshold). The percentage of reimbursed costs is termed the loss-sharing ratio. Across the Medicare PPSs, this ratio ranges from 50% to 80% (Table VIII.1). The inpatient psychiatric facility (IPF) payment system applies two loss-sharing ratios, which are intended to reflect a declining marginal per diem cost of furnishing inpatient psychiatric care over the course of a stay. In that system, the first loss-sharing ratio (80%) applies to eligible per diem costs up to and including the median number of covered days across all inpatient psychiatric facilities. The second ratio (60%) applies to costs associated with days in a stay beyond the median number of IPF covered days. The unit of payment in the IPF PPS is the day. However, its outlier policy parameters are calculated on a per stay basis and a provider's stay-level costs are examined relative to its stay-level payments in determining eligibility for outlier payments. This method eliminates an incentive to manipulate the timing of submitting charges, which would occur if outlier eligibility were determined using *actual* daily costs and charges. Alternatively, outlier policies can be developed based on daily costs *averaged* over a stay. Depending on patterns of average daily costs and average lengths of stay, stay-level outlier policies and average daily outlier policies can result in different distributions of outlier payments. Some suggest that the assessment of stay-level costs and payments is a better reflection of the overall financial impact of a case on a provider (CMS 2004).

Using a payment system's specified target percentage and loss-sharing ratio(s), each year CMS uses the latest available Medicare claims data and cost-to-charge ratios to calculate a payment system's loss amount for the coming payment year. This dollar amount is printed in the proposed and final rules in the *Federal Register*.

C. Data Sources

Our outlier policy analyses are performed using a 10% random sample of SNFs and all their associated stays in 2001. This file, which was used in the case-mix classification analyses presented in earlier chapters, was described in Chapter III.

Total per diem costs and NTA per diem costs were calculated from the SNF claims data and 2001 SNF cost reports. Cost-to-charge ratios were derived from cost reports and applied to charges on SNF claims, as described in detail in Chapter III. Total per diem payments were derived using the RUG-III payments associated with all claims in a stay divided by the number of covered days in a stay. Because NTA payments are not identified separately on the claims, we simulated expected NTA payments by estimating the proportion of total payments that are attributable to NTA services. These estimates were made for each RUG-III group.

Stays from facilities that had missing cost report data and that resulted in missing or clearly erroneous CCRs were excluded from the universe of stays from which the 10% data sample was selected. Other data quality flags were applied to the stays, as described in Chapter III. After applying these flags, the 10% sample includes 170,783 stays. After excluding another 530 stays (or about 0.33% of stays) due to missing RUG-III group assignments, the file used in the outlier policy analyses included 170,253 stays, furnished by a total of 1,386 SNFs.

D. Outlier Payment Policy Approaches and Parameters Analyzed for the SNF PPS

Budget-Neutral Policies

Budget-neutral outlier policies were simulated to accompany payment rates under the existing RUG-III case-mix classification system. That is, outlier payments and the RUG-III base rate reductions necessary to maintain budget neutrality were calculated and applied. Under both the total cost and NTA cost policy analyses, we reduced total RUG-III rates. Regarding NTA cost policies, this means that in order to finance NTA outlier compensation, we reduced *all* payments (i.e., routine, NTA, and therapy payments combined) rather than only reducing the estimated portion of RUG-III payments attributable to NTA services. An NTA policy could also be financed solely by reductions in the estimated NTA payments. The project's TAP was generally, though not unanimously, in favor of reducing all payments to finance either outlier approach (SNF TAP 2004). Reasons for reducing all payments to finance outliers include the point that, in any PPS, cross-subsidization by providers inherently occurs across their cost components and cases. Further, some analysts suggest that RUG-III payments provide overpayments for therapy costs relative to NTA costs and recommend more explicit cross-subsidization from SNF therapy payments to SNF NTA payments (MedPAC 2004).

Total Cost Outlier Policies

The total cost policies allow for outlier payments due to extraordinary NTA costs (such as high prescription drug costs), routine costs (such as in facilities with very high nurse staffing ratios), or therapy costs. Under a total cost outlier approach, a stay in a given RUG-III group is eligible for outlier compensation when its total costs pass a threshold relative to its full RUG-III group payment. Thus, the dollar level of the threshold varies by RUG-III group. Because outlier thresholds are set in relation to a payment system's case-mix classification groups, the extent of the variation in the thresholds in these outlier analyses is bounded by the variation in the payment weights of the RUG-III system. Stays with total costs exceeding their threshold receive compensation for a portion of their total costs exceeding this total threshold.

Five total cost outlier policy targets are presented in this chapter: outlier payments equal to 1%, 2%, 3% , 4 %, and 5% of total SNF PPS payments in 2001. For example, SNF PPS payments in 2001 were roughly $11 billion; thus, a 1% total outlier policy has a target of about $110 million. These target percentages encompass the range of target percentages used in all other Medicare PPSs except for in the long-term care hospital per case PPS, which has an 8% target. In selecting a target, other payment systems have selected the target that results in 10% of stays being eligible for outlier payments (e.g., the inpatient psychiatric facility PPS (CMS 2003, CMS 2004)). In this study, a 4% total cost target results in 10% of SNF stays receiving outlier payments.

NTA Cost Outlier Policies

Under the NTA cost outlier approach, a stay in a given RUG-III classification group is eligible for outlier compensation only when its NTA costs pass a threshold relative to the estimated amount of that RUG-III group payment attributable to NTA. The dollar level of the threshold varies by RUG-III group, and the variation is bounded by the variation in the "routine" component of the RUG-III group. This is because in the RUG-III case-mix classification system, these payment weights are intended to account for variation in resource use associated with nursing care and NTA. Stays with NTA costs exceeding their threshold receive compensation for a portion of their NTA costs in excess of the threshold. Under our NTA cost policy analyses, a case may not receive NTA outlier payments if its total SNF PPS payments exceeded its total costs. Among cases not meeting this condition (i.e., those cases that are eligible for NTA outlier payments except for the fact that their RUG-III payments are greater than their total costs), RUG-III payments exceed total costs by a range of about 12% to 18%.

A total of eight NTA cost outlier policy targets were developed in this study. As with the total cost policies, 1% through 5% NTA targets were first developed. The estimated share of SNF payments attributable to NTA in 2001 is about 16%, thus a 1% NTA outlier policy has a target of about $18.2 million. None of these five targets, however, results in 10% of SNF stays receiving NTA outlier payments. Because this share has been used as a guideline in selecting outlier policies, we also derived an NTA target that meets the 10% share guideline. In this study, this share equals a 15% target. Because of the extent of the mismatch between RUG-III payments and NTA resource use and the desire among some stakeholders to redistribute all NTA payments differently, we also wanted to simulate an NTA target that equals the entire share of SNF PPS payments attributable to estimated NTA payments. However the resulting target— 16%— is very close to the target that captures 10% of SNF stays (the 15% target). In the tables below, we present three targets: the 15% target because it meets the 10% share guideline, and the 5% target because it is the target closest to the 10% share guideline from among the group of targets used in most of Medicare’s other payment systems (the 1% through 5% group). We also present a 10% target, because it shows a middle-range alternative between the 5% and 15% targets. Results for all eight NTA cost outlier policies can been seen in Appendix VIII.

Loss-Share Approaches

In both the total cost and NTA cost policies, two loss-sharing ratios are applied, reflecting declining marginal costs of care over the course of a SNF stay. A loss-sharing ratio of 80% is applied to eligible per diem costs from day one of a stay to days up to and including the median number of covered days. A case's eligible per diem costs equal its total eligible costs divided by the number of covered days.[25] A loss-sharing ratio of 60% is applied to eligible per diem costs associated with any days in excess of the median. As noted above, the use of two loss-sharing ratios, and their application to costs based on a provider group's median number of covered days, is applied in Medicare's other per diem PPS, the inpatient psychiatric facility PPS.

Across all SNF stays, the median number of covered days is eighteen. Mean and median lengths of SNF stays vary markedly across provider groups, and are correlated in particular with a facility's hospital-based or free-standing status. Applying an overall median value in determining the loss-share percentage applied to eligible outlier costs results in the eligible outlier costs of shorter-length stays being compensated at an overall higher rate than that of longer-length stays. To offer two policy approaches to this issue, we applied two alternative sets of medians in calculating outlier payments under each total cost policy target and each NTA cost policy. Under the first approach, a common median (18 days) was calculated across all SNF stays and applied in outlier payment determinations. Under the second approach, separate medians for hospital-based and free-standing facilities were calculated (11 days and 21 days,

[25] In determining eligible outlier costs, fixed-loss amounts were wage-adjusted to account for area variation in provider costs. This adjustment is consistent with outlier policies in other Medicare PPSs.

respectively), and applied in outlier payment determinations for these two facility groups.[26] Results from the common median approach are presented in Sections E through G. Summary statistics of the separate median approach are presented as well, and detailed findings of the separate median approach are presented in terms of percent change (e.g., of average outlier payment amounts) between the common median and separate median approaches.

Using per Stay Versus per Diem Costs in Determining Outlier Payments

The SNF outlier payment policies developed here analyze payments per stay relative to costs per stay. This approach was used because some policymakers view the assessment of stay-level costs and payments as a better reflection of the overall financial impact of a case on a provider than daily or average daily costs and payments (CMS 2004), and because it follows the policy decision regarding the inpatient psychiatric facility per diem PPS (CMS 2003, CMS 2004). However, sound outlier policies can be developed based on daily costs averaged over the course of a stay. (*Actual* daily cost policies are undesirable because they create incentives to manipulate the timing of charges claimed across days within a stay). Among SNF stays, per diem costs are higher on average among hospital-based stays; however, total lengths of stay are lower among these stays. Consequently, per stay costs are similar on average between hospital-based and free-standing stays. Because of these patterns, SNF outlier policies based on average daily costs result in higher outlier payments in general for hospital-based facilities relative to free-standing ones. Because per stay payments and costs vary less by hospital-based/free-

[26] Similarly, one could establish a cut-off between shorter and longer SNF stays, allocate facilities to these 2 groups (regardless of hospital-based/free-standing status), calculate medians from these groups, and apply them in calculating outlier payments for cases in these groups.

standing status, however, per stay outlier policies are somewhat more neutral with respect to hospital-based/free-standing status. Preferences between these approaches may depend in part on one's view of the rationality of the differing lengths of stays and costs among these facility types. Appendices VIII.1 through VIII.4 show the distribution of SNF total and NTA costs per stay and per day.

E. Total Cost Outlier Policy Findings

Table VIII.2 shows summary statistics from the total cost outlier policy simulations, using the "common median" loss-share calculation. For reference, the top panel of the table first shows average RUG-III payments per stay, average costs per stay, and the ratio of RUG-III payments to costs, calculated across *all* SNF stays (rather than only those stays receiving outlier payments). The rest of the top panel shows summary statistics regarding the outlier policies, again calculated across all SNF stays. The bottom panel shows the summary statistics calculated *only* across stays actually receiving outlier payments.

Table VIII.2: Summary Statistics under Five Total Outlier Payment Policy Simulations

	No Outlier Policy	1%	2%	3%	4%	5%
All stays:						
Mean RUG-III payments per stay[a]	6,763					
Mean costs per stay	7,083					
Ratio of RUG-III payments to costs[a]	0.95					
Fixed-loss amount	n.a.	11,091	7,067	5,255	4,162	3,408
Percent stays with outlier payments	n.a.	1.5	3.7	6.4	9.3	12.5
Percent of total payments that are outlier payments[b]	n.a.	1.0	2.0	2.9	3.8	4.8
Ratio of total payments to costs[b]	n.a.	0.95	0.95	0.95	0.95	0.95
Mean outlier payment per stay	n.a.	67	133	197	260	323
Stays with outlier payments:						
Mean RUG-III payments per stay[a]	n.a.	11,500	10,070	9,187	8,588	8,135
Mean costs per stay	n.a.	28,465	21,720	18,314	16,184	14,634
Ratio of RUG-III payments to costs[a]	n.a.	0.40	0.46	0.50	0.53	0.56
	n.a.					
Mean total payments per stay[b]	n.a.	15,867	13,434	12,002	11,062	10,339
Ratio of total payments to costs[b]	n.a.	0.56	0.62	0.66	0.68	0.71
Mean outlier payments per stay	n.a.	4,481	3,562	3,083	2,805	2,593

[a] "RUG-III" payments are existing PPS payments.
[b] "Total" payments are RUG-III payments plus outlier payments.

Summary Statistics

As seen on the top panel, under the five total cost outlier policy targets developed, fixed-loss amounts range from an estimated $11,091 (1% policy) to $3,408 (5% policy), meaning that a SNF receives outlier payments for stays with losses in excess of these values. As expected given the steps involved in determining fixed-loss amounts, the overall percent of total payments paid as outlier payments is equal to or near the target percentages. The percent of stays receiving outlier payments is higher than the percent of total payments paid out as outlier payments. The differences between these two sets of percentages is greater than the differences found in early outlier policy simulations under the acute hospital PPS (e.g., ProPAC 1987) and the inpatient rehabilitation facility PPS (e.g., Carter, et al. 2002). This comparison with other provider groups indicates that the SNF total cost distribution is somewhat more skewed than the acute hospital and inpatient rehabilitation facility total cost distributions. Under the five SNF outlier policies, outlier payments range from an average across all stays of $67 (1% policy) to $323 (5% policy).

Under each of the five total cost outlier policies, outlier payments are paid for stays that otherwise would incur substantial losses. This is seen in the second panel of Table VIII.2, which shows average RUG-III payments and costs of these stays and ratios of their RUG-III payments to costs. For example, under a 1% policy, the average outlier-eligible stay costs $28,465 but receives only $11,500 in RUG-III payments absent an outlier policy payment.[27] Under the policies discussed here, the RUG-III payment-to-cost ratios range from 0.4 (1% policy) to 0.56 (5% policy) and ratios of total payments to costs improve substantially, ranging from 0.56 (1% policy) to 0.71 (5% policy). As the outlier target percentage increases, average outlier payment per outlier declines. This is because the higher targets are identifying relatively less costly stays as outliers. Table VIII.3 is a corollary to Table VIII.2. It shows the summary statistics from the outlier policy simulations in which separate median covered day values are applied to hospital-based and free-standing stays when determining outlier reimbursements for outlier-eligible stays. The results in these two tables are very similar. This is expected, because both present means of the simulations across all stays, rather than by RUG-III group or by facility characteristics. The key differences to note are the lower fixed-loss amounts that are identified and applied in the second methodology, and the lower average outlier payments made to those stays with outliers.

[27] Note that RUG-III payments absent an outlier policy and outlier payments do not sum to total payments under an outlier policy. This is because of the reduction in RUG-III payments applied in the simulations in order to fund the outlier policy.

The lower average payments reflect the application of lower thresholds and the impact of the lower average losses among those stays now eligible for outlier payments under the second methodology. Impacts of the second methodology are discussed more fully below, in the findings of stays grouped by their facility characteristics.

Table VIII.3: Summary Statistics under Five Total Outlier Payment Policy Simulations, (Using Alternative Loss-Share Methodology)

	No Outlier Policy	1%	2%	3%	4%	5%
All stays:						
Mean RUG-III payments per stay[a]	6,763					
Mean costs per stay	7,083					
Ratio of RUG-III payments to costs[a]	0.95					
Fixed-loss amount	n.a.	10,868	6,898	5,099	4,017	3,291
Percent stays with outlier payments	n.a.	1.6	3.9	6.7	9.8	13.0
Percent of total payments that are outlier payments[b]	n.a.	1.0	2.0	2.9	3.9	4.8
Ratio of total payments to costs[b]	n.a.	0.95	0.95	0.95	0.95	0.95
Mean outlier payment per stay	n.a.	67	133	197	261	322
Stays with outlier payments:						
Mean RUG-III payments per stay[a]	n.a.	11,425	9,967	9,096	8,495	8,061
Mean costs per stay	n.a.	28,087	21,379	18,015	15,890	14,417
Ratio of RUG-III payments to costs[a]	n.a.	0.41	0.47	0.50	0.53	0.56
Mean total payments per stay[b]	n.a.	15,614	13,167	11,768	10,832	10,157
Ratio of total payments to costs[b]	n.a.	0.56	0.62	0.65	0.68	0.70
Mean outlier payments per stay	n.a.	4,303	3,397	2,937	2,664	2,480

Note: This table reflects outlier payment policy simulations that apply two separate median covered day values (by hospital-based and free-standing status) when determining whether eligible per diem outlier costs are reimbursed at 80% of costs (for days up to the median) or at 60% of costs (for days beyond the median).

[a] "RUG-III" payments are existing PPS payments.
[b] "Total" payments are RUG-III payments plus outlier payments.

Outlier Stay and Payment Percentages, by RUG-III Category

Table VIII.4 shows the percent of cases receiving outlier payments and the percent of total payments paid as outlier payments under the five outlier policies, by RUG-III category. The incidence of outliers is highest in the special care category—ranging from 1.9% of all stays (1% policy) to 12.6% of stays (5% policy). Under most of the five policies, the outlier payment percentage is close to the outlier stay percentage. This pattern diverges in the rehabilitation category under the 3%, 4%, and 5% policies, where the outlier stay percentage is almost three times higher than the outlier payment percentage. Under the 5% policy, for example, an estimated 10.6% of rehabilitation category stays must receive outlier payments in order to spend 3.8% of total payments on outliers. This suggests that as the pool of available outlier funds increases and the outlier eligibility threshold falls, the average loss on rehabilitation category stays is much smaller than the average loss on stays in most of the other categories. This pattern is also found to some degree in two RUG-III categories with very low rates of Medicare-covered stays—impaired cognition and behavior problems only.

Table VIII.4: Outlier Stay and Payment Percentages under Five Total Outlier Payment Policies, by RUG-III Category

	Outlier Stay Percents					Outlier Payment Percents				
RUG-III Category	**1%**	**2%**	**3%**	**4%**	**5%**	**1%**	**2%**	**3%**	**4%**	**5%**
Rehabilitation	0.6	2.1	4.3	7.2	10.6	0.5	1.1	1.9	2.8	3.8
Extensive Services	1.3	3.0	4.7	6.7	8.6	1.8	3.6	5.3	6.8	8.3
Special Care	1.9	4.2	6.8	9.5	12.6	2.1	4.1	5.9	7.7	9.5
Clinically Complex	0.8	2.4	4.0	5.2	7.4	1.1	2.2	3.2	4.3	5.2
Impaired Cognition	0.0	0.0	0.6	0.8	2.4	0.0	0.0	0.2	0.5	0.9
Behavioral Problems	0.0	0.0	0.0	2.2	4.3	0.0	0.0	0.0	0.9	2.0
Reduced Physical Function	0.3	1.6	2.2	3.1	4.3	0.4	1.6	3.0	4.3	5.6

We also examined these percentages by the forty-four RUG-III groups (see Appendix Exhibit VIII-6). Under each total cost outlier policy, the RUG-III groups with the highest outlier incidences are generally rehabilitation low/B (RLB), special care/A (SSA), and extensive services/3 (SE3). Under the five policies the outlier stay percentages in RLB range from 4.8% to 24.5%, and range from roughly 2% to over 10% for both SSA and SE3. Relative to other rehabilitation groups, high incidence in a rehabilitation low group makes sense clinically, given the inverse relationship found between patients' need for or ability to withstand therapy and their need for NTA services, and given the less controllable and predictable nature of NTA costs compared with therapy costs.

Among the fourteen rehabilitation groups, the outlier payment percentages are consistently lower than the outlier stay percentages. However among the RUG-III groups in the next three categories (extensive services, special care, and clinically complex), the outlier payment percentages are often higher than the stay percentages. This indicates that the average loss on SNF stays in the rehabilitation groups is smaller than the average loss on stays in many of the other RUG-III groups.

Outlier Stay and Payment Percentages, by Facility Characteristic

Tables VIII.5 through VIII.9 analyze the outlier policy simulations in terms of characteristics of the SNFs in which stays occur. Stays are assessed in terms of six broad facility characteristics: free-standing or hospital-based status; rural or urban status; ownership and control status (non-profit, government, for-profit/chain-owned, or for-profit/ individually-owned); facility size (0-49 Medicare-certified beds, 50-99 Medicare beds, and 100+ Medicare beds); Medicare volume (9% or fewer patient days that are Medicare-covered days; >9%-15%

Medicare days; >15%-25%; and >25%); and the nine Census divisions. Facility size and Medicare volume groupings were identified based on examinations of the distribution of these variables in the 2001 data. (Appendix Table VIII.5 lists states by the nine Census divisions).

Table VIII.5 shows the outlier stay percentage and outlier payment percentage by these characteristics under the common median approach. Outlier stays and payments are not distributed evenly among various types of facilities. In particular, hospital-based SNFs, small SNFs, and high volume SNFs (>25% Medicare days) have higher percentages of outlier stays and payments than free-standing SNFs, larger facilities, and lower volume ones. As our project's earlier descriptive analyses found, hospital-based SNFs account for a disproportionate share of these facilities. (Our earlier analyses showed that hospital-based facilities account for 25% of SNFs with nineteen or fewer Medicare-certified beds and for 77% of SNFs with 25% or more Medicare days). West South Central SNFs have higher percentages than other Census divisions, followed by facilities in the West North Central and East South Central divisions. Mid-Atlantic SNFs generally have the lowest percentages.

Table VIII.5: Outlier Stay and Payment Percentages under Five Total Outlier Payment Policies, by Facility Characteristics

	Outlier Stay Percents					Outlier Payment Percents				
Characteristic	**1%**	**2%**	**3%**	**4%**	**5%**	**1%**	**2%**	**3%**	**4%**	**5%**
All	*1.5*	*3.7*	*6.4*	*9.3*	*12.5*	*1.0*	*2.0*	*2.9*	*3.8*	*4.8*
Free Standing	0.6	1.6	2.7	4.0	5.6	0.4	0.7	1.1	1.4	1.8
Hospital Based	4.2	10.1	17.2	24.7	32.7	3.9	7.8	11.3	14.5	17.5
Rural	1.8	4.3	7.2	10.4	13.7	1.3	2.4	3.5	4.6	5.6
Urban	1.4	3.6	6.2	9.0	12.1	0.9	1.9	2.8	3.7	4.6
Non-profit	2.6	6.4	10.7	15.5	20.8	2.0	3.9	5.8	7.5	9.2
Government	3.6	9.1	15.6	22.4	29.6	2.7	5.8	8.7	11.4	13.9
For-profit	0.7	1.7	3.0	4.5	6.1	0.5	0.9	1.3	1.7	2.1
Chain	0.6	1.6	2.9	4.4	6.0	0.5	0.8	1.2	1.6	2.0
Non-chain	0.9	2.2	3.6	4.9	6.6	0.5	1.0	1.5	2.0	2.4
0-49 Medicare beds	2.6	6.4	10.7	15.4	20.5	2.0	4.0	5.8	7.6	9.3
50-99 Medicare beds	0.7	1.9	3.6	5.3	7.3	0.4	0.9	1.3	1.9	2.4
100+ Medicare beds	0.6	1.6	2.7	4.0	5.5	0.4	0.7	1.1	1.4	1.8
0%-9% Medicare days	0.8	2.1	3.5	5.1	6.8	0.5	1.0	1.5	2.0	2.5
10%-15% Medicare days	0.6	1.5	2.6	4.0	5.6	0.4	0.7	1.0	1.4	1.7
16%-25% Medicare days	0.3	0.8	1.6	2.6	3.8	0.2	0.3	0.5	0.7	1.0
>25% Medicare days	3.7	8.9	15.1	21.7	28.8	3.3	6.4	9.3	11.9	14.5
New England	1.0	2.3	4.1	6.1	8.4	0.6	1.2	1.8	2.4	3.0
Mid-Atlantic	0.7	2.1	4.1	6.4	9.3	0.7	1.1	1.6	2.2	2.8
South Atlantic	1.4	3.7	6.1	8.8	11.8	0.8	1.8	2.7	3.5	4.4
East North Central	1.3	3.2	5.8	8.6	11.8	0.8	1.6	2.5	3.3	4.2
East South Central	1.6	4.1	7.0	10.8	15.1	1.2	2.3	3.4	4.5	5.6
West North Central	2.2	5.3	9.2	13.0	17.0	1.8	3.5	5.1	6.6	8.1
West South Central	4.1	9.3	14.4	19.7	24.5	3.0	5.7	8.1	10.1	12.0
Pacific	1.1	3.0	5.0	7.0	9.1	0.7	1.6	2.4	3.2	4.0
Mountain	1.3	3.1	5.1	7.2	9.9	0.9	1.8	2.6	3.4	4.2

Table VIII.6 is the corollary to this table, showing results from the alternative loss-share calculation method. Under this approach, medians for hospital-based and free-standing facilities were calculated separately (11 days and 21 days, respectively), and then applied in outlier payment determinations for these two facility types. (Under the common median approach seen on Table VIII.5, a single median (18 days) was calculated across all SNF stays and applied in the outlier payment determinations). As was seen on Table VII.3, fixed-loss amounts are lower when using the separate median approach than the common median approach. Lower fixed-loss amounts allow a greater number of stays to be eligible for outlier compensation.

Table VIII.6: Outlier Stay and Payment Percentages under Five Total Outlier Payment Policies, by Facility Characteristics (Using the "Separate Median" Loss-Share Methodology)

	Outlier Stay Percents					Outlier Payment Percents				
Characteristic	**1%**	**2%**	**3%**	**4%**	**5%**	**1%**	**2%**	**3%**	**4%**	**5%**
All	*1.7*	*4.4*	*7.6*	*10.8*	*14.2*	*1.0*	*2.0*	*2.9*	*3.8*	*4.8*
Free Standing	0.7	1.8	3.2	4.7	6.4	0.5	0.9	1.3	1.7	2.1
Hospital Based	4.8	12.0	20.3	28.8	37.3	3.6	7.2	10.4	13.5	16.3
Rural	2.1	5.1	8.4	12.0	15.5	1.2	2.4	3.5	4.6	5.6
Urban	1.6	4.2	7.3	10.5	13.9	0.9	1.9	2.8	3.7	4.6
Non-profit	3.0	7.6	12.7	18.1	23.8	1.9	3.7	5.5	7.2	8.9
Government	4.3	10.6	18.3	26.0	33.6	2.6	5.5	8.2	10.8	13.2
For-profit	0.8	2.0	3.6	5.2	7.0	0.5	1.0	1.4	1.9	2.3
Chain	0.7	1.9	3.5	5.1	6.9	0.5	0.9	1.3	1.8	2.2
Non-chain	1.0	2.4	3.8	5.4	7.1	0.6	1.2	1.8	2.3	2.8
0-49 Medicare beds	3.0	7.5	12.7	18.0	23.4	2.0	3.9	5.7	7.4	9.0
50-99 Medicare beds	0.8	2.3	4.3	6.3	8.4	0.4	0.9	1.4	2.0	2.5
100+ Medicare beds	0.7	1.8	3.2	4.7	6.3	0.4	0.8	1.2	1.6	2.0
0%-9% Medicare days	0.9	2.4	4.1	5.8	7.7	0.5	1.1	1.7	2.3	2.8
10%-15% Medicare days	0.7	1.8	3.2	4.7	6.5	0.4	0.8	1.2	1.6	2.0
16%-25% Medicare days	0.3	1.0	2.0	3.1	4.5	0.2	0.4	0.6	0.8	1.1
>25% Medicare days	4.2	10.6	17.9	25.3	32.8	3.1	6.0	8.7	11.3	13.6
New England	1.2	2.7	4.9	7.2	9.9	0.6	1.3	1.9	2.5	3.1
Mid-Atlantic	0.8	2.7	5.1	7.7	10.8	0.7	1.2	1.8	2.4	3.0
South Atlantic	1.7	4.3	7.2	10.2	13.5	0.8	1.8	2.7	3.5	4.4
East North Central	1.5	3.8	6.9	10.1	13.5	0.8	1.7	2.6	3.5	4.4
East South Central	1.9	4.8	8.4	12.9	17.2	1.1	2.2	3.2	4.3	5.4
West North Central	2.6	6.4	10.7	15.2	19.3	1.7	3.3	4.9	6.4	7.8
West South Central	4.7	10.7	16.5	22.1	27.2	2.9	5.5	7.8	9.8	11.6
Pacific	1.5	3.5	5.8	8.5	11.3	0.9	1.8	2.6	3.4	4.2
Mountain	1.3	3.5	5.8	8.1	10.5	0.7	1.5	2.3	3.1	3.8

Note: The appendix tables reflect outlier policy simulations that apply two separate median covered day values (by hospital-based and free-standing status) when determining whether eligible per diem outlier costs are reimbursed at 80% of costs (for days up to the median) or at 60% of costs (for days beyond the median)

Comparing Tables VIII.5 and VIII.6 indicates that the separate median outlier policy approach reaches a greater number of SNF stays both overall and by each facility type. For example, under the 5% outlier policy, 14.2% of all stays receive outlier payments under the separate median approach compared to 12.5% under the common median approach. Among free-standing facility stays, 6.4% of stays receive outlier payments under the separate median approach compared to 5.6% under the other approach. Among hospital-based facility stays, 37.3% of stays receive outlier payments under the separate median approach versus 32.7% under the other approach. Again, the lower fixed-loss amount enables these outlier stay shares to be greater under the separate median approach.

In terms of the percent of total (RUG-III plus outlier) payments paid out as outlier compensation, a comparison of Tables VIII.5 and VIII.6 shows that these percentages are higher among free-standing facilities under the separate median approach, and are lower among hospital-based facilities under the separate median approach. This means that under the separate median approach (compared to the common median approach), average outlier payments are higher for free-standing facility stays and are lower for hospital-based facility stays. In essence, the lower outlier payments made to hospital-based facilities under the separate median approach create opportunities to provide outlier compensation to a greater share of both free-standing and hospital-based stays. This pattern is also seen by two facility characteristics highly correlated with free-standing facilities— propriety facilities and facilities with lower Medicare volume.

Outlier Payments per Stay, by Facility Characteristic

Table VIII.7 shows the average outlier payment amount calculated across all stays under the five total cost policies, by facility characteristics, under the common median approach. Overall, average outlier payments under these policies range from $67 (1% policy) to $323 (5% policy). Average outlier payments differ by facility type. In particular, markedly higher average payments are made to stays in SNFs that are hospital-based, non-profit, government-owned, small in terms of bed number, have high Medicare volume shares, and are in the West South Central states. As noted above, a disproportionate share of facilities across these characteristics are hospital-based. Under a 1% outlier policy, average outlier payments for stays in SNFs with these characteristics are $112-$177. Under a 5% policy, average payments in these SNFs are $545-$878. Among stays in SNFs that are free-standing, for-profit, larger in size, or have lower Medicare volume, average payments are about $30-$77 under the 1% policy and about $144- $458 under the 5% policy.

Table VIII.7: Average Outlier Payments Across all Stays under Five Total Outlier Payment Policies, by Facility Characteristics

Characteristic	Average Outlier Payments per Stay				
	1%	2%	3%	4%	5%
All	*67*	*133*	*197*	*260*	*323*
Free-standing	30	55	81	107	134
Hospital Based	177	361	537	709	878
Rural	77	148	215	279	342
Urban	64	129	192	255	318
Non-profit	112	226	337	444	551
Government	137	301	456	608	758
For-profit	35	64	93	122	152
Chain	34	62	90	118	148
Non-chain	37	74	110	143	175
0-49 Medicare beds	115	228	336	441	545
50 -99 Medicare beds	31	63	97	133	169
100+ Medicare beds	30	58	85	112	140
0% -9% Medicare days	31	65	99	132	166
10%-15% Medicare days	29	53	79	104	131
16%-25% Medicare days	15	28	43	59	77
>25% Medicare days	164	326	481	633	781
New England	48	93	137	181	228
Mid-Atlantic	49	83	122	165	211
South Atlantic	54	118	179	238	296
East North Central	53	108	164	222	282
East South Central	73	141	208	277	349
West North Central	100	195	286	373	458
West South Central	172	338	485	616	738
Pacific	52	114	174	233	289
Mountain	56	115	168	219	269

Table VIII.8 is the corollary to Table VIII.7, and shows average outlier payments (calculated across all stays) under the separate median approach. Comparing Tables VIII.7 and VIII.8 shows that average outlier payments are higher for free-standing facility stays under the separate median approach (e.g., $159 per stay, 5% policy) than under the common median approach ($134 per stay, 5% policy), and payments are slightly lower for hospital-based stays ($804 per stay, 5% policy) under the separate median approach compared to the common median approach ($878 per stay, 5% policy). Average payments also are slightly higher among stays in proprietary facilities and those with lower Medicare volume shares. Across the remaining characteristics, average outlier payments are very similar under the two loss-share methodologies.

Table VIII.8: Average Outlier Payments Across all Stays under Five Total Outlier Payment Policies, by Facility Characteristics (Using the "Separate Median" Loss-Share Methodology)

	Average Outlier Payments per Stay				
Characteristic	**1%**	**2%**	**3%**	**4%**	**5%**
All	*67*	*133*	*197*	*260*	*323*
Free Standing	35	66	97	128	159
Hospital Based	163	329	491	649	804
Rural	75	147	214	279	342
Urban	65	129	193	255	317
Non-profit	108	216	322	424	526
Government	130	282	430	574	715
For-profit	38	71	104	137	169
Chain	37	68	99	130	162
Non-chain	46	91	133	169	205
0 -49 Medicare beds	113	223	328	430	530
50-99 Medicare beds	32	66	103	140	178
100+ Medicare beds	32	62	92	121	151
0%-9% Medicare days	35	76	115	153	190
10%-15% Medicare days	33	64	93	123	154
16%-25% Medicare days	16	32	49	68	88
>25% Medicare days	156	305	450	591	729
New England	48	96	142	189	237
Mid-Atlantic	54	91	132	176	223
South Atlantic	53	118	179	238	296
East North Central	57	114	173	232	292
East South Central	70	134	199	267	338
West North Central	96	184	273	359	440
West South Central	167	325	467	594	709
Pacific	51	109	167	221	274
Mountain	56	116	170	221	273

Note: The appendix tables reflect outlier policy simulations that apply two separate median covered day values (by hospital-based and free-standing status) when determining whether eligible per diem outlier costs are reimbursed at 80% of costs (for days up to the median) or at 60% of costs (for days beyond the median).

Ratio of Total Payments to Costs, by Facility Characteristic

Table VIII.9 shows ratios of total payments to costs per stay under the five outlier policies, by facility characteristic, using the common median method of determining outlier payment amounts. Total payments include the outlier compensation made to eligible stays plus the regular RUG-III payments, and account for budget-neutrality reductions made to the base PPS payment rates. The base rate reductions apply to all SNF stays, and are necessary for financing the outlier payments. For comparison, ratios of RUG-III payments to costs are shown. The RUG-III payment-to-cost ratios refer to payments made assuming no outlier policy is applied, and thus do not include any outlier payment or base PPS payment rate reductions.

Table VIII.9: Ratios of Total Payments to Costs under Five Total Outlier Payment Policies by Facility Characteristics

	Ratio of RUG-III Payments to Costs per Stay	*Ratio of Total Payments to Costs per Stay*				
Characteristic	**No Outlier Policy**					
		1%	**2%**	**3%**	**4%**	**5%**
All	*0.95*	*0.95*	*0.95*	*0.95*	*0.95*	*0.95*
Free-standing	1.08	1.08	1.07	1.06	1.06	1.05
Hospital Based	0.59	0.61	0.63	0.65	0.67	0.68
Rural	0.94	0.94	0.94	0.94	0.94	0.95
Urban	0.96	0.96	0.96	0.96	0.96	0.96
Non-profit	0.78	0.79	0.79	0.80	0.81	0.82
Government	0.68	0.69	0.71	0.72	0.74	0.75
For-profit	1.09	1.08	1.07	1.07	1.06	1.06
Chain	1.09	1.08	1.08	1.07	1.07	1.06
Non-chain	1.07	1.06	1.06	1.05	1.05	1.04
0-49 Medicare beds	0.80	0.81	0.81	0.82	0.83	0.84
50 -99 Medicare beds	1.05	1.05	1.04	1.04	1.03	1.03
100+ Medicare beds	1.09	1.08	1.08	1.07	1.06	1.06
0% -9% Medicare days	1.06	1.05	1.05	1.04	1.04	1.03
10%-15% Medicare days	1.09	1.08	1.08	1.07	1.06	1.06
16%-25% Medicare days	1.12	1.11	1.10	1.09	1.08	1.08
>25% Medicare days	0.66	0.67	0.69	0.70	0.72	0.73
New England	1.02	1.02	1.01	1.01	1.01	1.00
Mid-Atlantic	0.99	0.99	0.98	0.98	0.98	0.97
South Atlantic	0.97	0.96	0.96	0.96	0.96	0.96
East North Central	0.95	0.95	0.95	0.95	0.95	0.95
East South Central	0.93	0.93	0.93	0.93	0.93	0.93
West North Central	0.86	0.86	0.87	0.88	0.88	0.89
West South Central	0.80	0.82	0.84	0.85	0.86	0.87
Pacific	1.02	1.02	1.02	1.02	1.02	1.01
Mountain	0.97	0.97	0.97	0.97	0.97	0.96

Overall, the average payment-to-cost ratio (PCR) remains 0.95, regardless of the application of an outlier policy. Average PCRs vary by facility characteristic, however, when outlier policies are applied. The largest positive impact on average PCRs is seen for hospital-based SNFs, non-profit or government facilities, high Medicare volume SNFs, and facilities in the West South Central states. The average PCR for hospital-based SNFs rises from 0.59 under no outlier policy to 0.68 under a 5% policy. For high-volume SNFs, average PCRs rise from 0.66 under no outlier policy to 0.73 under the 5% policy. These findings are consistent with the higher average outlier payment amounts seen in Table VIII.7 for stays in these facilities. In contrast, average PCRs remain positive but decline somewhat for facilities that are free-standing, for-profit, large in terms of bed number, and have lower Medicare volume shares. Among these facilities, average PCRs decline from between 1.07 and 1.12 under no outlier policy to between 1.03 and 1.08 under a 5% outlier policy. This indicates that for these facilities, reductions made to the base PPS payment rates to permit budget neutrality have a greater (and negative) impact than the outlier payments themselves.

Changes in average PCRs that occur when applying the separate median method rather than the common median method are not shown, because the changes are negligible. The largest impact of the separate median method is seen for the PCR of hospital-based SNFs. This ratio falls under the 5% outlier policy from 0.68 under the common median method to 0.67 under the separate median method (not shown).

F. NTA Cost Outlier Policy Findings

In this section, the eight NTA cost policies are analyzed and compared. Comparisons across NTA cost policies alone (or across total cost policies alone) are particularly useful if one approach is clearly more desirable in principle over the other. The section following this (Section G) highlights two particular total cost and NTA cost policies that appear to be the most analogous to each other, compares them, and suggests tentative overall conclusions regarding outlier policies under the RUG-III SNF PPS. Comparisons within an outlier policy approach (i.e., within only total cost or only NTA cost policies) and across these approaches are useful as well, if policymakers are neutral in principle with respect to the two approaches. As noted above, the total cost outlier policies are designed based on percents of *total* SNF costs, while the NTA cost policies are designed based on percent of costs *only* for NTA services. Under our NTA cost policies, NTA outlier compensation is conditional on a stay's total costs exceeding its total PPS payments. In our data, among NTA stays not meeting this condition, RUG-III payments exceed total costs by a range of 12% to 18%.

Tables VIII.10 through VIII.17 present the same analyses of NTA cost outlier policies as are seen in Tables VIII.2 through VIII.9 above for total cost policies, except that mainly three NTA outlier targets are analyzed. The 15% NTA target is presented because it meets the policy guideline used in some payment systems of providing outlier compensation to 10% of cases. The 5% target is presented because it is the target closest to the "10% share" guideline from among the group of lower target percentages seen in Medicare's other PPSs. The 10% target is presented because it shows a mid-range policy between the 5% and 15% targets. (Appendix Exhibits VIII.7 through VIII.13 show results for each of the eight NTA outlier policies simulated).

Summary Statistics

Under the eight NTA cost outlier policy targets developed, fixed-loss amounts range from a high of $10,956 in NTA costs (1% policy) to a low of $1,096 (16% policy), meaning that a SNF receives NTA outlier payments for stays with losses in excess of these levels of NTA costs. As expected given the steps involved in determining fixed-loss amounts, the overall percent of total payments paid as outlier payments is equal to or at least near the target percentages. Unlike in the total cost policies (Table VIII.2), however, the percent of stays receiving NTA outlier payments is lower than the percent of NTA payments distributed as NTA outlier payments. This reflects the fact that the SNF NTA cost distribution is more skewed than the SNF total cost distribution. Under the eight NTA outlier policies, outlier payments range from an average across all SNF stays of $11 (1% policy) to $153 (16% policy).

Table VIII.10: Summary Statistics under Eight NTA Outlier Payment Policy Simulations

	No Outlier Policy	1%	2%	3%	4%	5%	10%	15%	16%
All stays:									
Mean RUG-III payments per stay[a]	6,763								
Mean costs per stay	7,083								
Ratio of RUG-III payments to costs[a]	0.95								
Fixed-loss amount	n.a.	10,956	7,168	5,512	4,477	3,760	1,948	1,177	1,096
Percent stays with outlier payments[b]	n.a.	0.3	0.7	1.2	1.7	2.3	5.8	10.0	10.7
Percent of total payments that are outlier payments[b]	n.a.	1.0	2.0	2.9	3.9	4.8	9.2	13.2	13.8
Ratio of total payments to costs	n.a.	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
Mean outlier payment per stay	n.a.	11	22	32	42	53	101	146	153
Stays with outlier payments:									
Mean RUG-III payments per stay[a]	n.a.	14,703	12,519	11,496	10,844	10,230	8,410	7,354	7,224
Mean costs per stay	n.a.	34,703	26,995	23,534	21,606	19,929	15,199	12,670	12,387
Ratio of RUG-III payments to costs[a]	n.a.	0.42	0.46	0.49	0.50	0.51	0.55	0.58	0.58
Mean total payments per stay[b]	n.a.	18,968	15,516	14,074	13,231	12,410	10,018	8,646	8,482
Ratio of total payments to costs[b]	n.a.	0.55	0.57	0.60	0.61	0.62	0.66	0.68	0.68
Mean outlier payments per stay	n.a.	4,292	3,041	2,638	2,460	2,264	1,739	1,454	1,425

[a] "RUG-III" payments are existing PPS payments.
[b] "Total" payments are RUG-III payments plus outlier payments.

The second panel of the table shows that NTA outlier payments substantially improve the ratio of payments to costs compared with ratios under a SNF PPS without outlier payments. Absent any NTA outlier policy (i.e., absent any outlier payments or corresponding base-rate reductions), RUG-III payment-to-cost ratios for the stays that would receive outlier payments under the eight NTA policies range from 0.42 (1% policy) to 0.58 (16% policy). This increase indicates that costs are much higher, relative to RUG-III payments, among the very top of the outlier distribution (stays affected by a 1% policy), than among stays further down the outlier distribution. With the NTA outlier payments, payment-to-cost ratios increase and range from 0.55 (1% policy) to 0.68 (16% policy). Payment-to-cost ratios under the total cost policies (Table VIII.2) were lower than these ratios absent the outlier payments, and higher than these ratios with outlier payments. Finally, the last row of the table shows that, as expected, average outlier payments per outlier case declines as the target increases. Despite the much smaller levels of funds available for outlier compensation in NTA cost policies than total cost policies, the average outlier payments are in the same general range across the two policy approaches. This feature, along with the relatively small percentages of stays receiving NTA outlier payments, is a reflection of the relatively skewed distribution of SNF NTA costs.

Table VIII.11 is a corollary to Table VIII.10, showing summary statistics for NTA outlier payments derived using the "separate median" method for hospital-based versus free-standing SNFs when calculating outlier compensation. Comparing Tables VIII.10 and VIII.11 suggests the two loss-share approaches result in less of a difference regarding NTA cost policies than regarding the total cost policies (Tables VIII.2 and VIII.3). For example, the NTA fixed-loss amounts and average NTA outlier payments per stay are roughly comparable under the two loss-share approaches. The separate median approach results in slightly lower NTA fixed-loss amounts ($3,717 versus $3,760 under a 5% target) and lower average outlier payments made to those stays with outliers ($2,229 versus $2,264 under a 5% target). The lower average payments reflect the application of lower thresholds and the impact of the lower average losses among these stays now eligible for outlier payments under the second methodology.

Table VIII.11: Summary Statistics under Eight NTA Outlier Payment Policy Simulations (Using the "Separate Median" Loss-Share Methodology)

	No Outlier Policy	1%	2%	3%	4%	5%	10%	15%	16%
All stays:									
Mean RUG-III payments per stay[a]	6,763								
Mean costs per stay	7,083								
Ratio of RUG-III payments to costs[a]	0.95								
Fixed-loss amount	n.a.	10,922	7,135	5,464	4,436	3,717	1,921	1,170	1,062
Percent stays with outlier payments[b]	n.a.	0.3	0.7	1.2	1.8	2.4	5.9	10.1	11.0
Percent of total payments that are outlier payments[b]	n.a.	1.0	2.0	2.9	3.9	4.8	9.1	13.0	13.8
Ratio of total payments to costs	n.a.	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
Mean outlier payment per stay	n.a.	11	22	32	42	53	100	144	153
Stays with outlier payments:									
Mean RUG-III payments per stay[a]	n.a.	14,643	12,523	11,480	10,798	10,200	8,379	7,339	7,169
Mean costs per stay	n.a.	34,491	26,979	23,466	21,497	19,846	15,115	12,641	12,261
Ratio of RUG-III payments to costs[a]	n.a.	0.42	0.46	0.49	0.50	0.51	0.55	0.58	0.58
Mean total payments per stay[b]	n.a.	18,840	15,504	14,025	13,144	12,345	9,950	8,606	8,390
Ratio of total payments to costs[b]	n.a.	0.55	0.57	0.60	0.61	0.62	0.66	0.68	0.68
Mean outlier payments per stay	n.a.	4,224	3,025	2,604	2,418	2,229	1,699	1,426	1,387

Note: This table reflects outlier payment policy simulations that apply two separate median covered day values (by hospital-based and free-standing status) when determining whether eligible per diem outlier costs are reimbursed at 80% of costs (for days up to the median) or at 60% of costs (for days beyond the median).

[a] "RUG-III" payments are existing PPS payments.
[b] "Total" payments are RUG-III payments plus outlier payments.

Outlier Stay and Payment Percentages, by RUG-III Category

Table VIII.12 shows, by RUG-III category, the percent of stays receiving NTA outlier payments and the percent of NTA payments paid out as NTA outlier payments under the three NTA cost policies we highlight. Across all three policies, the incidence of NTA outlier-compensated stays is highest in the special care category, ranging from 4.3% of stays (5% policy) to 17.6% of stays (15% policy). Under the total cost policies discussed above, the share of payments paid out as outlier compensation is generally in the range of the share of stays receiving outlier compensation (e.g., see Table VIII.4). These percentages diverge under NTA cost policies, reflecting the more skewed nature of NTA costs and the fact that a smaller share of stays deplete the NTA target. This pattern of stay percentages being smaller than payment percentages under NTA policies is consistent across the RUG-III categories.[28]

Table VIII.12: Outlier Stay and Payment Percentages under Three NTA Outlier Payment Policies, by RUG-III Category

	Outlier Stay Percents			Outlier Payment Percents		
RUG-III Category	**5%**	**10%**	**15%**	**5%**	**10%**	**15%**
Rehabilitation	1.2	3.9	7.9	2.5	6.4	11.1
Extensive Services	3.0	7.2	12.5	10.4	20.6	30.9
Special Care	4.3	10.4	17.6	17.5	33.9	50.0
Clinically Complex	1.9	5.0	10.1	8.8	18.2	28.4
Impaired Cognition	0.0	1.4	4.7	0.0	4.6	13.1
Behavioral Problems	0.0	2.2	4.3	0.0	2.4	15.8
Reduced Physical Function	0.2	1.6	5.5	0.6	7.7	20.9

Note: Based on RUG-III assignments that reflect 75 percent or more of a stay. This restriction was used to allow outlier results associated with a RUG-III group to reflect stays that are generally characterized by that RUG-III assignment. Results for eight RUG-III groups are based on extremely small (n<30) sample sizes: IB2, IA2, BB1, PE2, PD2, PC2, PB2, PA2.

An examination of the outlier stay percentages by the forty-four RUG-III groups (Appendix Exhibit VIII.8) indicates that the rehabilitation low-B (RLB) group consistently incurs the highest incidence of NTA outlier-compensated cases. This pattern is consistent with prior analyses under this project and by other research suggesting that stays with minimal rehabilitation costs and high NTA costs are classified by providers in the low and medium rehabilitation groups, in order to receive the highest possible RUG-III payments for these cases.

[28] Under the total cost policies (from Table VIII.4), this pattern diverges somewhat in the rehabilitation category. Total cost policy outlier stay percentages for the simulations of 2%-5% are 2-3 times *higher* than the outlier payment percentages. As the pool of available total cost outlier funds increased and the outlier-eligibility threshold fell, the average total loss on rehabilitation category stays was much smaller than the average total loss on stays in most of other RUG-III categories.

In contrast, under the total cost policies relatively high incidence was seen in three RUG-III groups: RLB, special care-A (SSA), and extensive services-3 (SE3). As the NTA outlier target increases, the NTA outlier stay incidences in the extensive services and special care groups begin to approach the incidence seen in RLB.

Outlier Stay and Payment Percentages, by Facility Characteristic

Tables VIII.13 through VIII.17 show the impact of the NTA cost outlier policies by facility characteristic. Table VIII.13 shows the percent of stays receiving outlier payments and the percent of total payments paid out as outlier compensation under the common median approach. As was seen in the total cost outlier policy findings as well, NTA outlier stays and payments are not distributed evenly across facilities. For example, SNFs that are hospital-based, small in terms of bed number, and have high Medicare volume have higher percentages of outlier stays and payments than free-standing SNFs, larger facilities, and lower Medicare volume ones. Also as seen under the total cost policies, SNFs in the West South Central and East South Central regions have the highest percentages of NTA outlier stays and payments. In many cases, the patterns by facility characteristic become more pronounced under the larger NTA outlier targets.

Table VIII.13: Outlier Stay and Payment Percentages under Three NTA Outlier Payment Policies, by Facility Characteristics

	Outlier Stay Percents			Outlier Payment Percents		
Characteristic	**5%**	**10%**	**15%**	**5%**	**10%**	**15%**
All	*2.3*	*5.8*	*10.0*	*4.8*	*9.2*	*13.2*
Free Standing	1.7	4.0	6.7	3.4	6.2	8.9
Hospital Based	4.2	11.3	19.8	11.2	21.4	29.8
Rural	2.6	6.3	10.9	5.7	10.4	14.7
Urban	2.3	5.7	9.8	4.6	8.9	12.9
Non-profit	2.9	7.6	13.5	6.5	12.8	18.5
Government	3.6	9.7	17.2	8.1	16.3	23.4
For-profit	1.9	4.4	7.4	3.8	7.0	9.9
Chain	1.9	4.6	7.6	3.9	7.2	10.2
Non-chain	1.6	3.8	6.3	3.2	6.0	8.6
0-49 Medicare beds	3.0	7.8	13.7	6.7	13.1	19.0
50-99 Medicare beds	2.0	4.8	8.0	4.3	7.8	11.1
100+ Medicare beds	1.6	3.9	6.6	3.1	5.9	8.5
0%-9% Medicare days	1.7	3.9	6.8	4.1	7.1	9.9
10%-15% Medicare days	1.7	4.0	6.9	3.0	5.8	8.5
16%-25% Medicare days	1.7	4.0	6.8	3.5	6.3	8.8
>25% Medicare days	3.8	10.1	17.7	9.2	18.2	25.8
New England	1.3	3.4	6.5	2.2	4.7	7.5
Mid-Atlantic	1.8	4.6	8.4	3.4	6.7	9.9
South Atlantic	2.4	5.8	9.8	4.3	8.7	12.7
East North Central	2.1	5.1	8.9	4.8	8.8	12.5
East South Central	2.1	5.4	9.7	6.0	10.1	14.0
West North Central	2.5	6.3	11.3	5.9	11.3	16.2
West South Central	4.8	11.8	19.2	10.0	18.5	25.4
Pacific	2.3	5.9	9.9	5.0	9.7	14.2
Mountain	2.3	6.0	10.2	4.2	8.9	13.3

Table VIII.14 is the corollary to Table VIII.13, showing outlier stay and payment percentages under the separate median approach to determining outlier payments. Under each of the total cost policies, the separate median approach had resulted in higher shares of stays receiving payments and generally in slightly lower shares of total payments paid out as outlier compensation (Table VIII.6). Among the NTA cost policies, however, the outlier stay and payment percentages vary much less between the two loss-share approaches.

Table VIII.14: Outlier Stay and Payment Percentages under Three NTA Outlier Payment Policies, by Facility Characteristics (Using the "Separate Median" Loss-Share Methodology)

	Outlier Stay Percents			Outlier Payment Percents		
Characteristic	**5%**	**10%**	**15%**	**5%**	**10%**	**15%**
All	*2.4*	*5.9*	*10.1*	4.8	9.1	13.0
Free-Standing	1.7	4.0	6.7	3.4	6.4	9.0
Hospital Based	4.3	11.4	19.9	10.8	20.6	28.7
Rural	2.6	6.4	10.9	5.6	10.3	14.4
Urban	2.3	5.8	9.9	4.6	8.8	12.7
Non-profit	3.0	7.7	13.6	6.3	12.5	18.0
Government	3.6	9.8	17.3	7.8	15.7	22.5
For-profit	1.9	4.5	7.4	3.9	7.1	10.0
Chain	1.9	4.6	7.6	4.0	7.3	10.2
Non-chain	1.6	3.9	6.4	3.2	6.1	8.7
0 -49 Medicare beds	3.0	7.9	13.7	6.6	12.9	18.4
50 -99 Medicare beds	2.1	4.9	8.1	4.4	7.9	11.1
100+ Medicare beds	1.6	3.9	6.6	3.2	6.0	8.5
0% -9% Medicare days	1.7	4.0	6.8	4.1	7.2	10.0
10%-15% Medicare days	1.7	4.1	6.9	3.1	5.9	8.6
16%-25% Medicare days	1.8	4.1	6.9	3.6	6.4	8.9
>25% Medicare days	3.9	10.3	17.8	8.9	17.6	24.9
New England	1.3	3.4	6.6	2.2	4.8	7.5
Mid-Atlantic	1.8	4.7	8.4	3.5	6.8	9.9
South Atlantic	2.5	5.9	9.8	4.4	8.7	12.6
East North Central	2.1	5.2	8.9	4.8	8.8	12.4
East South Central	2.1	5.4	9.8	5.9	9.9	13.6
West North Central	2.5	6.5	11.3	5.9	11.1	15.8
West South Central	4.9	12.0	19.2	9.8	18.1	24.7
Pacific	2.3	6.0	10.2	5.0	9.6	13.9
Mountain	2.3	6.1	10.0	4.2	8.9	13.2

Thus, the findings suggest that in terms of the percent of stays receiving outlier compensation, the separate median approach is worth considering if a total cost outlier policy is preferred. However the separate median approach make little difference in the case of NTA cost policies.

Outlier Payments per Stay, by Facility Characteristic

Table VIII.15 illustrates average NTA outlier payment amounts, calculated across all stays under the NTA cost outlier policies, under the common median approach and by facility characteristic. Average NTA outlier payments are over twice as large for hospital-based facilities relative to free-standing ones. Average NTA payments are 1.2 to 1.8 times higher for facilities that are non-profit, government-owned, small in terms of bed number, high Medicare volume share, and in the West South Central states. As noted earlier, a disproportionate share of facilities across these characteristics are hospital-based.

Table VIII.15: Average Outlier Payments Across all Stays under Three NTA Outlier Payment Policies, by Facility Characteristics

Characteristic	Average Outlier Payments per Stay 5%	10%	15%
All	*53*	*101*	*146*
Free-standing	41	74	104
Hospital Based	87	180	268
Rural	57	105	148
Urban	51	100	145
Non-profit	62	125	185
Government	69	145	218
For-profit	46	83	117
Chain	48	86	121
Non-chain	37	69	97
0-49 Medicare beds	63	127	187
50 -99 Medicare beds	51	92	128
100+ Medicare beds	39	73	103
0% -9% Medicare days	46	79	109
10%-15% Medicare days	37	71	101
16%-25% Medicare days	46	81	113
>25% Medicare days	77	160	239
New England	26	56	87
Mid-Atlantic	43	82	120
South Atlantic	47	95	137
East North Central	52	96	135
East South Central	61	103	142
West North Central	54	104	150
West South Central	99	193	276
Pacific	59	116	170
Mountain	42	91	136

Table VIII.16 is the corollary to Table VIII.15, and shows average NTA outlier payments under the separate median approach. As was seen under the total cost policies as well, these NTA cost policy findings show a pattern of higher outlier payments to SNFs that are free-standing, proprietary, and have lower Medicare volume shares, and lower outlier payments for hospital-based facilities under the separate median approach when compared with the common median approach. However, the differences resulting from the two loss-sharing approaches are much smaller, and in some cases negligible, under the NTA cost policies.

Table VIII.16: Average Outlier Payments Across all Stays under Three NTA Outlier Payment Policies, by Facility Characteristics (Using the "Separate Median" Loss-Share Methodology)

Characteristic	Average Outlier Payments per Stay		
	5%	10%	15%
All	*53*	*100*	*144*
Free-standing	42	76	106
Hospital Based	84	172	255
Rural	56	104	146
Urban	51	100	143
Non-profit	60	122	179
Government	66	139	207
For-profit	47	85	118
Chain	49	88	122
Non-chain	29	71	99
0-49 Medicare beds	62	124	181
50 -99 Medicare beds	52	93	128
100+ Medicare beds	40	74	104
0%-9% Medicare days	47	81	110
10%-15% Medicare days	38	72	103
16%-25% Medicare days	47	83	114
>25% Medicare days	74	154	228
New England	27	57	87
Mid-Atlantic	43	83	120
South Atlantic	47	95	136
East North Central	53	96	134
East South Central	60	101	139
West North Central	53	101	145
West South Central	96	187	266
Pacific	59	115	167
Mountain	43	91	135

Ratio of Total Payments to Costs, by Facility Characteristic

Table VIII.17 shows average ratios of *total* payments to *total* costs per stay under the NTA cost policies, using the common median approach. We show ratios of total payments to total costs, since ratios of a component of payments and costs are not as relevant for assessing the overall impact of outlier policies on facilities. For reference, the first column of the table shows ratios of RUG-III payments to total costs (*i.e.*, in which no outlier policy is applied). Under the total cost policies (Table VIII.9), the findings had shown that outlier policies increase the payment-to-cost ratios for hospital-based facilities and actually decrease them for free-standing facilities. As this table shows, under the NTA cost policies the payment-to-cost ratio increases for hospital-based facilities but does not decrease for free-standing ones.

Table VIII.17: Ratios of Total Payments to Costs under Three NTA Outlier Payment Policies, by Facility Characteristics

Characteristic	Ratio of RUG-III Payments to Costs per Stay	Ratio of Total NTA Payments to Costs per Stay		
	No Outlier Policy	5%	10%	15%
All	*0.95*	*0.95*	*0.95*	*0.95*
Free-standing	1.08	1.08	1.08	1.08
Hospital Based	0.59	0.60	0.61	0.62
Rural	0.94	0.94	0.94	0.94
Urban	0.96	0.96	0.96	0.96
Non-profit	0.78	0.78	0.78	0.79
Government	0.68	0.68	0.69	0.69
For-profit	1.09	1.08	1.08	1.08
Chain	1.09	1.09	1.09	1.08
Non-chain	1.07	1.06	1.06	1.06
0-49 Medicare beds	0.80	0.80	0.80	0.81
50 -99 Medicare beds	1.05	1.05	1.05	1.05
100+ Medicare beds	1.09	1.09	1.08	1.08
0%-9% Medicare days	1.06	1.06	1.05	1.05
10%-15% Medicare days	1.09	1.09	1.08	1.08
16%-25% Medicare days	1.12	1.12	1.11	1.11
>25% Medicare days	0.66	0.66	0.67	0.68
New England	1.02	1.02	1.01	1.01
Mid-Atlantic	0.99	0.99	0.99	0.99
South Atlantic	0.97	0.97	0.97	0.97
East North Central	0.95	0.95	0.95	0.95
East South Central	0.93	0.93	0.93	0.93
West North Central	0.86	0.86	0.86	0.86
West South Central	0.80	0.81	0.82	0.82
Pacific	1.02	1.02	1.02	1.02
Mountain	0.97	0.97	0.97	0.97

When comparing the payment-to-cost ratios of a total cost policy and an NTA cost policy of roughly comparable target sizes— the 3% total cost policy (Table VIII.9) and the 15% NTA cost policy — the findings indicate that the total cost policy improves the ratios more so than the NTA cost policy for facilities that are hospital-based, non-profit, government-owned, small in terms of bed number, and have high Medicare volume shares. The NTA policy improves the ratios relatively more (or reduces the positive ratios less) for free-standing SNFs, for-profit facilities, larger facilities, and lower Medicare volume SNFs.

G. Additional Comparisons between Total and NTA Policy Approaches

This final section of the chapter presents additional comparisons between the total cost outlier approach and the NTA cost outlier approach, and offers overall comments and explanations regarding the two approaches. We first describe the "overlap" of the number of stays receiving outlier compensation under the two approaches. In a more detailed illustration, we next compare total and NTA policies that each compensate 10% of SNF stays.

Overlap of Stays Receiving Outlier Compensation under Total and NTA Policies

Table VIII.18 shows the percent of SNF stays that receive outlier compensation under each of the three NTA cost policies that also receive outlier compensation under each of the five total policy simulations.

Table VIII.18: Percent of Stays Receiving Outlier Payments under Three NTA Policies that also Receive Outlier Payments under Total Policies

Total Cost Outlier Policies	NTA Cost Outlier Policies		
	5%	10%	15%
	Percent of NTA Outlier Stays also Receiving Total Outlier Payments		
1%	34	18	11
2%	58	38	26
3%	71	54	40
4%	79	65	52
5%	85	73	62

As expected, among stays receiving outlier payments under a given NTA cost policy, the share receiving payments under the total cost policy approach increases as the total cost policy targets increase. For example, in the 5% NTA cost policy this share ranges from 340% of stays that receive outlier payments under a 1% total cost policy to 85% under the 5% total cost policy. This occurs simply because the amount of outlier funds available under the total cost policies increase when moving from the 1% to 5% total cost policy simulations.

More importantly, one might expect to see larger "overlaps" than Table VIII.18 indicates. That is, one might expect that larger shares of NTA cost outlier recipients would also be total cost outlier recipients. This might be expected particularly since NTA cost policies are so much smaller in aggregate dollars than the total cost policies (for example, a 1% total cost target equals roughly $110 million, while a 1% NTA cost target equals only about $18.2 million). However, the table suggests that a substantial share of SNF stays have total costs that meet the total cost outlier criteria but have NTA costs that do not meet the NTA cost outlier criteria.

One factor that may partly explain the level of overlap seen on Table VIII.18 is that facility-specific routine (including skilled nursing) costs are used, by necessity, in total cost outlier policy determinations. Thus, all SNF stays treated in facilities with very high average routine costs are more likely to meet a total cost policy outlier criteria, regardless of a stay's actual level of routine, therapy, and NTA costs, than stays in facilities with more moderate routine costs. Very high facility-specific routine costs may reflect several factors, including a severely ill case mix, cost inefficiencies, or variations in cost accounting practices. GAO (2002b) found that these factors explain the higher routine costs of hospital-based SNFs relative to free-standing ones. Further, the facility-level routine costs associated with Medicare stays may be understated among free-standing SNFs, given the difficulty of identifying routine costs for Medicare patients in facilities with large shares of stays covered by multiple payers. The unavailability of stay-level routine costs and the uncertainty of the contribution of different causes of routine cost variation are critical challenges in developing routine payment PPS weights using the cost-report approach. Similarly, these features reflect a weakness that exists in a total cost outlier policy approach but that does not exist in an NTA cost outlier approach.

Comparing Total and NTA Policies that Each Compensate 10% of SNF Stays

To further explore the characteristics of stays that do and do not overlap, in terms of receiving outlier compensation under the two approaches, the 4% total cost and 15% NTA cost policies are compared. These are selected because each provides outlier compensation to about 10% of SNF stays, a share that, as noted, has been used as a guideline in selecting specific outlier targets in other Medicare PPSs. As seen on Table VIII.18, 52% of stays receiving outlier compensation under the 15% NTA cost policy simulations also receive outlier payments under the 4% total cost policy.

To help describe the stays affected by either of these policies, Table VIII.19 shows the average payments, costs, and payment-to-cost ratios by cost component of the stays. Payments in this table reflect payments prior to receiving any outlier payment or base PPS payment rate reduction. The RUG-III payment-to-cost ratio is highest among the stays receiving payments only under the NTA cost policy. This may suggest that overall, the NTA cost policy approach compensates stays that do not need outlier compensation as severely (i.e., that have lower costs relative to their RUG-III payments) as stays compensated under the total cost policy approach.

Table VIII.19: Average Costs and Payments per Stay of Stays Receiving Outlier Payments under the 4% Total and 15% NTA Outlier Policies

	Number stays (in sample)	RUG-III Payments	Total Costs	Ratio of RUG-III Payments to Total Costs	NTA Payments	NTA Costs	Ratio of Estimated NTA Payments to NTA Costs	Therapy and Routine Payments	Therapy and Routine Costs	Ratio of Estimated Therapy and Routine Payments to Costs
Total and NTA Policies	8,805	8,217	16,735	0.49	1,393	5,220	0.27	6,824	11,515	0.59
Total Policy Only	6,994	9,055	15,490	0.58	1,450	1,325	1.09	7,605	14,165	0.54
NTA Policy Only	8,280	6,437	8,347	0.77	1,051	3,174	0.33	5,385	5,172	1.04

Note: Data are from the "common median" approach described in the text. RUG-III payments are payments prior to any outlier payments.

Table VIII.20 shows the distribution of stays under the 4% total cost and 15% NTA cost policies by RUG-III category. In the stays that overlap these two outlier policies, almost 70% are classified in the rehabilitation category, but significant shares are also in the extensive services and special care categories (16.8% and 10.8%, respectively). A similar distribution is found among the remaining stays in the “NTA cost policy only” group. In contrast, the majority of the stays in the “total cost policy only” group are classified in the rehabilitation category (89.5 percent).

Table VIII.20: RUG-III Category of Stays Receiving Outlier Payments under a 4% Total Outlier Policy and a 15% NTA Outlier Policy

RUG-III Category	Distribution of Stays across RUG-III Categories among Stays in:		
	Total and NTA Policies	Total Policy Only	NTA Policy Only
Rehabilitation	69.2	89.5	65.0
Extensive Services	16.8	4.3	17.3
Special Care	10.8	3.3	11.6
Clinically Complex	2.8	2.1	4.8
Impaired Cognition	0.1	0.0	0.4
Behavioral Problems	0.0	0.1	0.1
Reduced Physical Function	0.3	0.6	0.9
Total	*100*	*100*	*100*

Table VIII.21 shows the distribution of stays under the two policies by facility characteristics. Among the stays that overlap the two outlier policies, 71.6% are hospital-based. The literature indicates that hospital-based facilities are oriented to SNF patients with extensive NTA service needs, and thus one might expect that the share of stays in the "NTA cost policy only" column would be high as well. Instead, only 27.4% of the stays in the NTA cost policy only subset are hospital based compared with 62.8% of stays in the "total cost policy only" subset. Thus, some combination of patient-specific therapy costs and facility-specific routine and skilled nursing costs is resulting in an emphasis of the total cost policy approach on hospital-based stays over free-standing stays. With respect to other facility characteristics, the table indicates that non-profit SNFs are heavily represented in the overlap subset and the total cost policy only subset (about 55% of stays are in non-profit facilities). For-profit chain facilities are most common among the NTA cost policy only stays (51.1%).

Small bed and low Medicare volume facilities are the most common in the overlap and total cost policy only subsets (generally 70% or higher).

Table VIII.21: Facility Characteristics of Stays Receiving Outlier Payments under the 4% Total and 15% NTA Outlier Policies

	Percent of Stays Receiving Outlier Payments		
Characteristic	Total and NTA Policies	Total Policy Only	NTA Policy Only
Free Standing	28.4	37.2	72.6
Hospital Based	71.6	62.8	27.4
Rural	23.2	24.0	22.5
Urban	76.9	76.0	77.5
Non-profit	55.5	55.1	32.8
Government	15.3	16.1	6.7
For-profit Chain	25.2	21.4	51.1
For-profit Non-Chain	4.1	7.4	9.4
0 -49 Medicare beds	72.7	71.5	44.9
50 -99 Medicare beds	15.3	13.3	25.4
100+ Medicare beds	12.0	15.2	29.8
0%-9% Medicare days	12.1	17.8	24.4
10%-15% Medicare days	10.0	13.7	27.9
16%-25% Medicare days	5.5	3.4	17.3
>25% Medicare days	72.5	65.1	30.4
New England	3.4	6.5	6.1
Mid-Atlantic	11.3	12.4	17.3
South Atlantic	16.6	18.8	19.6
East North Central	17.5	20.1	18.3
East South Central	8.7	8.7	5.7
West North Central	10.2	10.4	6.2
West South Central	18.8	13.6	10.6
Pacific	8.3	5.7	11.0
Mountain	4.3	2.9	5.1

H. Concluding Comments

An outlier policy is an important payment policy tool that defrays some of the costs providers bear for cases whose treatment costs are extraordinarily high relative to their regular PPS payments. Each of Medicare's other prospective payment systems includes some type of a total cost outlier policy. The main advantage of a total cost policy is that it allows for compensation of any type of extraordinary patient cost, whether routine or ancillary. The main disadvantage of a total cost policy in the SNF setting is that SNF routine costs are not available at the case level, and this introduces a bias in estimating a case's total costs and its eligibility for outlier compensation under a total cost policy. Because of this problem with SNF total cost outlier policies, and because of a focus on improving payments particularly for SNF NTA costs, we developed several NTA cost outlier policies, as well as several total cost outlier policies.

The findings presented in this chapter show that the incidence of outlier cases, under both the total cost and NTA cost outlier approaches, is highest in the RUG-III special care category. The greatest number of outlier stays is found in the rehabilitation category (namely because roughly 85% of SNF stays are classified in this category). Across the forty-four RUG-III groups, the incidence of outlier cases is the highest in the rehabilitation low/B (RLB), special care/A (SSA), and extensive services/3 (SE3) groups.

The findings show that outlier policies improve the average payment-to-cost ratio or PCR of hospital-based SNFs in particular, and of facilities that are non-profit, government-owned, and have high shares of Medicare volume. PCRs decline for facilities that are free-standing, for-profit, large in terms of bed number, and have lower shares of Medicare volume. The decreased

PCRs indicate that these facilities are disadvantaged by the PPS base rate reductions necessary to fund the outlier policies more so than these facilities are aided by outlier compensation.

In comparing the results of the total cost policies with the NTA cost policies, hospital-based stays in particular fare better under total cost policies and free-standing stays in particular fare better under NTA cost policies. For example, hospital-based SNF PCRs increase more under total cost policies than they increase under comparable NTA cost policies, and the free-standing SNF PCRs fall less under NTA cost policies than they fall under comparable total cost policies. One reason the total cost policies may favor hospital-based SNFs is that they have substantially higher average routine costs (although the causes and relative importance of the causes of these higher costs are not well understood (GAO 2002b)). Further, the availability of facility-level routine costs, rather than case-level costs, creates an upward bias in the total cost estimates of all cases treated in facilities with high facility-level average costs.

We also analyzed two alternatives for determining the amount of outlier compensation for stays eligible for the compensation— a "common median" and "separate median" method. Applying an overall median length of stay in the outlier compensation calculations results in the eligible outlier costs of shorter-length stays being compensated at an overall higher rate than that of longer-length stays. In contrast, the separate median method accounts for the large difference seen in the median lengths of stay of hospital-based versus free-standing facilities in the outlier compensation calculations. In the total cost policies, the separate median approach results in a higher share of stays (across all facility characteristics) receiving outlier compensation under the common median approach. It also results in larger outlier payments for stays particularly in free-standing SNFs than under the common median approach. These higher shares of outlier stays, and larger average outlier payments to free-standing SNF stays, basically occur because of the lower fixed-loss amounts and are funded by a reduction in the average outlier payments for hospital-based SNF stays seen in the separate median approach. This overall pattern is apparent when comparing the common median and separate median approaches under the NTA cost policies as well, however the size of the impact is much smaller, and in terms of many facility characteristics the impact is negligible. This suggests that the two methodologies are worth considering if a total cost policy is preferred over an NTA cost policy.

In conclusion, this chapter presented the impact of several total cost and NTA cost policies by RUG-III category and group and by facility characteristics, and discussed issues that may explain the findings. With this information, policymakers will be better able to make decisions regarding the value of incorporating an outlier policy into the SNF PPS, the merits and relative impacts of total cost policies versus NTA cost policies, the relative effects of various sizes of outlier targets, and the impact of using alternative methodologies in the outlier compensation calculations that account for some differences between hospital-based and free-standing facilities.

IX. Implications for Research and Policy

Findings from this study have implications for both research and policy. The extensive analyses we have conducted with multiple data sources shed light on strengths and weaknesses of the available data. Despite some limitations in the data, the models we developed provide "building blocks" that are helpful in considering options for improving Medicare payment for SNFs. In this chapter, we highlight research and policy implications flowing from those building blocks.

A. Research

MDS Variables on Service Use

One of our more important findings recognized that the manner in which the MDS currently asks service use questions allows the indicator to refer to 'use of services' to the SNF stay or the prior hospital stay. Although this construct is not technically a problem in determining the health status of a patient, it dilutes the service use indicator's effectiveness as a predictor of NTA costs during the SNF stay. To isolate use of specific services during the SNF stay, we interacted the MDS indicators with SNF claims' indicators for the same services. We were able to do this because our analysis file included MDS and corresponding claims data for the same stay. Beyond specific research projects, however, these two sources of data are not routinely merged. The implication of this interaction illustrates that consideration should be given to modifying the MDS to account for use of particular services, specifically during SNF stays.

Hospital and SNF Diagnoses on Claims

Although we originally focused on diagnosis information from SNF claims, we discovered through our research that the coding of diagnoses was extremely uneven across SNFs. In particular, while hospital-based SNFs tended to fully code diagnoses for their patients, freestanding SNFs were less diligent. Whereas patients of hospital-based SNFs appeared to have more secondary diagnoses than those of freestanding SNFs based on SNF claims, the difference in the number of secondary diagnoses literally disappeared when we compared the coding with that of their respective prior hospital stays. Using only SNF diagnosis information could have led, for example, to the erroneous conclusion that patients of hospital-based SNFs had more comorbidities than did patients of freestanding SNFs.

A related issue is that hospital-based SNFs often (23% of the time) used the rehabilitation procedure code (V57) to specify the primary diagnosis on SNF claims, whereas freestanding SNFs hardly ever (0.5% of the time) used this code for the primary diagnosis. Hence, this diagnosis from SNF claims acts, in part, as a proxy for hospital-based or freestanding status.

Because of the complications with the information on diagnoses from SNF claims, we chose to use diagnosis information from prior hospital stays in many of our analyses. Prior hospital stays' diagnoses were used, for example, in the New Profile and DRG-related analyses. If information on diagnoses is made integral to any new SNF payment model that is adopted, strong incentives will be provided to improve and unify the coding of diagnoses on SNF claims.

At this point, it would be appropriate to begin substituting SNF claims' diagnosis information for data from the prior hospital stays.

Disconnect between MDS Assessment and Claims Data

Claims information on ancillary service use and charges for SNF stays did not coincide with periods covered by MDS assessments, and most SNF stays involved multiple MDS assessments. In effect, cost information for ancillary services use referred only to the total duration of the SNF stay and not to particular days during the SNF stay, whereas MDS assessments refer to more specific periods during that stay. Where there were multiple MDS assessments for a given SNF stay, we weighted the scores on those assessments by the number of days during the stay covered by each assessment. While this approach represents our best attempt to relate patient characteristics (from the MDS) to charges (from the claims), more precise estimates could be derived if charge data were available to correspond with days covered by each MDS assessment. Consequently, we could not determine how costs might have changed over the course of the stay with the available data. Anecdotal observations by knowledgeable individuals assert that per diem costs tend to decline over the course of a SNF stay—but a valid methodology to measure such a trend is currently absent. To the extent that future research could begin to relate changes in costs over the duration of a stay, payment policies would need to incorporate this time varying occurrence.

Hospital-Based and Freestanding SNFs

The large difference in per diem costs between hospital-based and freestanding SNFs raised questions about whether the reasons behind the disparity warranted consideration for payment adjustment. Differences in case-mix and outcomes of care are examples of reasons why differential payments may be reasonable. Although our analysis of outlier payment options suggested that both NTA and nursing costs are contributors to higher overall per diem costs of hospital-based patients, the absence of outcome information elicits ambiguity on whether higher costs are necessary. On the other hand, less ambiguous differences in accounting procedures provide a clear example of why payment adjustments are not appropriate.

B. Policy

We developed models to address three aspects of the current SNF PPS that require attention:

(1) There is a general consensus that the RUG-III classification system does not adequately account for variations in NTA costs. To that end, we focused much of our efforts on developing models that improved the variance explanation of NTA costs, with three of our models developed to focus specifically on NTA costs. Although the DRG-related model was not developed specifically for NTA, it provides yet another option for adjusting Medicare payments for NTA services.

(2) The "fee-schedule" for rehabilitation therapy under the current SNF PPS promotes prodigal utilization of resources. We concentrated on developing one "need based" model for

rehabilitation therapy, in which amount of payment for rehabilitation therapy was determined primarily by patients' characteristics. The DRG-related model could also be considered as an alternative to the current rehabilitation therapy payment schedule.

(3) There is a need to update nursing payment weights over time to reflect changes in patient needs, service provision, and technology. Unlike for NTA and rehabilitation therapy services, information on nursing costs is routinely available only on a facility, rather than individual, basis. Thus, we were unable to develop person-level models for nursing care like those for NTA and rehabilitation therapy services. Only the DRG-related model, when applied to total costs, provides a means for using routinely collected information to update payment weights, including nursing services. The fact that the data requires us to assign the facility's average nursing cost to each patient in that facility to develop patient-level costs may be problematic, however, particularly since nursing costs tend to be a considerably higher percentage of total costs for SNFs than they are for other Medicare providers (e.g., acute care hospitals).

Our models on the various components of SNF services can serve as building blocks for ways to improve Medicare payment for SNFs. The relative merits of using the different models depend on the answers to these fundamental policy questions: (a) How much change is desirable? (b) Which service components should be addressed? (c) What tradeoffs are important in the selection of models?

How Much Change?

Not surprisingly, wide ranges of opinions exist concerning how much change to the SNF PPS is desirable. On the one hand, this system has been in effect for six-years, and participating SNF providers have adapted to its rules and incentives. Anecdotal information suggests that providers have learned to operate efficiently under the SNF PPS and would be reluctant to see major changes in its structure, including the RUG-III classification system underpinning the payment system (Maxwell 2003). Moreover, since many state Medicaid nursing home reimbursement systems also use a RUG-related classification system for adjusting case-mix, minimizing changes in the use of RUG-III would promote ease of operation for SNFs participating in both Medicare and Medicaid programs. On the other hand, the perceived problems with the current SNF PPS suggest that improvement in payment methodology is warranted. Whether improvements to the SNF PPS reflect minor RUG refinements or more wholesale changes will depend on how much weight is placed on the various criteria discussed above.

Which Components?

This question asks which components of the SNF PPS are likely to be the most profitable in terms of the SNF payment system. Most observers tend to rank payments for NTA services as the most important area for improvement, followed by rehabilitation therapy, and then by nursing.

Providers, consumer representatives, and policy analysts agree that RUG-III does not adequately adjust payments for NTA services. This perception has been supplemented by multiple research studies, all of which showed that RUG-III explains only about 5% of the variation in NTA costs. Costs for certain NTA services (e.g., IV therapy) may simply be much greater than expected NTA payments. The three NTA-specific models, as well as the DRG-based model, can be utilized to address the need for an improved payment strategy for NTA services. We reviewed policy related criteria of these models in Chapter VII.

The fee schedule nature in which the SNF PPS pays rehabilitation therapy services is another area to investigate for improving SNF payments. Some believe that fee schedule compensation has given providers too much latitude over payment for rehabilitation therapy services and provides strong incentives for providers to respond to financial, rather than clinical, incentives. The rehabilitation therapy-specific model, as well as the DRG-based model, can be evaluated relative to the need for an improvement in payment for rehabilitation therapy services.

The ability to update nursing payment weights using routinely collected data is a very desirable objective, given changes in patient needs, service provision, and technology. Stakeholders and policy analysts generally agree that this update should be done. Because administrative data on person-level nursing resource use has not been available, Medicare has had to rely on special staff time measurement studies of samples of patients to estimate relative nursing costs by patient characteristics. The RUG-III nursing weights were derived in this manner. There is a real dilemma here. Such studies are expensive and take a long time to conduct if sufficiently representative samples are to be obtained. But analytical efforts using routinely collected data on nursing costs confront the serious obstacle that information is available only on facility-level average nursing costs.

Choosing Among Different Models

When multiple models are available for consideration, choices among them depend on the weight one places on the different criteria used to compare them. For example, we have developed three models specifically focusing on NTA costs (four if we include the DRG-based model). In general, the more complex the model, the better its variance explanation, and the more clinically meaningful it is. At the same time, the more complex the model, the greater the data requirements, and the more complex the administrative procedures for implementing the classification system. Earlier in this report, we discussed the statistical and other criteria that we applied in assessing each of the models. Adoption of a "preferred model" for a given SNF service component will depend on how much weight policy makers place on the different criteria.

Unlike Medicare's PPS for other provider types, the SNF PPS currently does not include an outlier payment policy to address unusually high cost cases. We provided a number of illustrative examples in which we varied outlier target (e.g., pool), loss-sharing ratios, and the fixed-loss amount, and discussed different approaches for financing the outlier payments. An outlier payment policy may also be considered in conjunction with different classification models to improve Medicare payment for SNFs.

In sum, the building block models provide various options for potential improvements in the SNF PPS. One possibility establishes a new classification model explicitly for NTA services. A second possibility establishes a new NTA classification model and replaces the rehabilitation therapy component of RUG-III with the "need based" model for that service. This option could function independently or in conjunction with one of the other NTA models. A third possibility uses the DRG-based model for both NTA and rehabilitation therapy services.

In each of these scenarios, nursing payments would continue to be paid by the RUG-III nursing payment weights. Beyond the building-block strategy, DRGs could be considered for case-mix adjusting total costs. Finally, an outlier payment policy could be established for Medicare SNFs, either alone or in combination with any of the changes in the patient classification system.

References

Abt Associates. 2000. *Variation in Prescribed Medication and Other Non-Therapy Ancillary Costs in Skilled Nursing Facilities: Potential RUG-III Refinements.* Contract no. 500-96-0003/TO#7. Cambridge, MA: Abt Associates

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. *Classification and Regression Trees.* Belmont, CA: Wadsworth International Group.

Briesacher, B., B. Stuart, and J. Doshi. 2002. *Medication Use by Medicare Beneficiaries Living in Nursing Homes and Assisted Living Facilities*. Washington, DC: U.S. Department of Health and Human Services.

Buchanan, J.L., P.L. Andrews, S.M. Haley, S.M. Paddock, and A.M. Zaslavxky. 2004. "Evaluating the Planned Substitution of the Minimum Data Set-Post Acute Care for Use in the Rehabilitation Hospital Prospective Payment System." *Medical Care* 42(2): 155–63.

Carter, G.M., M.B. Buntin, and O. Hayden. 2002. *Analyses for the Initial Implementation of the Inpatient Rehabilitation Facility Prospective Payment System.* Report No. MR-1500-CMS. Santa Monica, CA: RAND Health.

Centers for Medicare and Medicaid Services (CMS). 2003a. "Change in Methodology for Determining Payment for Extraordinarily High-Cost Cases Cost Outliers under the Acute Care Hospital Inpatient and Long-Term Care Hospital Prospective Payment Systems." *Federal Register* 68(43, June 8): 10420.

———. 2003b. "Medicare Inpatient Psychiatric Facility Prospective Payment System, Proposed and Final Rules." *Federal Register* 68(229, November 28): 66919-66978.

———. 2003c. "Prospective Payment System for Inpatient Psychiatric Facilities; Proposed Rule." *Federal Register* 68(229, November 28): 66919-66978.

———. 2004a. "Prospective Payment System for Inpatient Psychiatric Facilities; Final Rule." *Federal Register* 69(219, November 15): 66921–67015.

———. 2004b. *SNF PPS Project Communication with CMS Project Officers and Advisors*. Contract no. 500-00-0025/TO#2. Baltimore, MD: Centers for Medicare and Medicaid Services.

Cotterill, P.G. 1986. "Testing a Diagnosis-Related Group Index for Skilled Nursing Facilities." *Health Care Financing Review* 7(4): 75–85.

Eilertsen, T.B., and DataPRO Project Investigators at the Colorado Foundation for Medical Care, the University of Colorado Health Sciences Center Division of Health Care Policy and

Research, Fu Associates Ltd., and Stepwise Systems Inc. 2002. *Rehabilitation Data Flag Analysis.* Baltimore, MD: Centers for Medicare and Medicaid Services.

Eilertsen, T.B., A.M. Kramer, R.E. Schlenker, and C.A. Hrincevich. 1998. “Application of Functional Independence Measure-Function Related Groups and Resource Utilization Groups-Version III Systems across Post-acute Settings.” *Medical Care* 36(5): 695–705.

Fries, B. 2003. Unpublished analyses of RUG-III on 1999 national claims data for Liu, K. in CMS Contract no. 500-00-0025/TO#2.

Fries, B., D. Schneider, W. Foley, M. Gavazzi, R. Burke, and E. Cornelius. 1994. “Refining a Case-Mix Measure for Nursing Homes: Resource Utilization Groups.” *Medical Care* 32(7): 668–85.

General Accounting Office (GAO). 1995. *Tighter Rules Needed to Curtail Overcharges for Therapy in Nursing Home.* GAO/HEHS-95-23. Washington, DC: GAO.

———. 2002a. *Federal Efforts to Monitor Resident Assessment Data Should Complement State Activities.* GAO-02-279. Washington, DC: GAO.

———. 2002b. *Medicare Payments Exceed Costs for Most but Not All Facilities*. GAO-03-183. Washington, DC: GAO.

Goldfield, I.N., E.C. McCullough, J.C. Vertrees, R.F. Averill, A.M. Tang, and J. Xiang. 2001. *Evaluation and Development of Patient Classification Systems for Long-Term Care Hospitals.* Report submitted to CMS. Baltimore, MD: Centers for Medicare and Medicaid Services.

Greene, W.H. 1997. *Econometric Analysis*. 3rd ed. Upper Saddle River, NJ: Prentice-Hall, Inc.

Hutt, E., M. Ecord, T.B. Eilertsen, E. Frederickson, J.C. Kowalsky, and A.M. Kramer. 2001. “Prospective Payment for Nursing Homes Increased Therapy Provision without Improving Community Discharge Rates.” *Journal of the American Geriatrics Society* 49 (8): 1071–79.

Institute for Health and Socio-Economic Policy. 2002. *US Hospitals: Utilizing or Exploiting Medicare Outliers?* Orinda, CA: Institute for Health and Socio-Economic Policy. Available at http://cna.igc.org/outliers/2002.

Keeler, E.B., G.M. Carter, and S. Trude. 1988. *Insurance Aspects of DRG Outlier Payments.* Report N-2762-HHS. Santa Monica, CA: The RAND Corporation.

Kramer, A.M., and DataPRO Project Investigators at the Colorado Foundation for Medical Care, the University of Colorado Health Sciences Center Division of Health Care Policy and Research, Fu Associates Ltd., and Stepwise Systems Inc. 2002. *Final Report*. Baltimore, MD: Centers for Medicare and Medicaid Services.

Kramer, A.M., T.B. Eilertsen, M.K. Ecord, and M.H. Morrison. 1999. *A Prospective Study of New Case-Mix Indices for Sub-Acute Care.* Report to National Subacute Care Association and the American Health Care Association. Denver, CO: Morrison Informatics and University of Colorado Health Sciences Center.

Lee, A.J., J.H. Huber, and W. B. Stason. 1997. "Factors Contributing to Practice Variation in Post-Stroke Rehabilitation." *Health Service Research* 32(2):197–221.

Levy, C.R., R. Fish, and A.M. Kramer. Forthcoming. "Site of Death in the Hospital versus Nursing Home among Medicare Skilled Nursing Facility Residents Admitted under Medicare's Part A Benefit." *Journal of the American Geriatrics Society*.

Liu, K., K. Black, and S. Maxwell. 2002. *Exploratory Analyses of Medicare Skilled Nursing Facility Costs in 1999.* Report to CMS. Contract no. 500-00-0025/TO#2. Baltimore, MD: Centers for Medicare and Medicaid Services.

Liu, K., and G. Kenney. 1993. "Impact of the Catastrophic Coverage Act and New Guidelines on Medicare Skilled Nursing Facility Use." *Inquiry* 30(1): 41–53.

Mahoney, F.I., and D.W. Barthel. 1965. "Functional Evaluation: The Barthel Index." *Maryland State Medical Journal* 14(2): 61–65.

Maxwell, S., J. Wiener, and K. Liu. 2003. *A Review of Issues in the Development and Implementation of the Skilled Nursing Facility Prospective Payment System*. Report to the Centers for Medicare and Medicaid Services (CMS), Contract no. 500-00-0025/TO#2. Baltimore, MD: Centers for Medicare and Medicaid Services.

Medicare Payment Advisory Commission (MedPAC). 1999. *Report to the Congress: A Framework for Considering Medicare Payment Policy Issues.* Washington, DC: MedPAC.

———. 2002. *Report to the Congress: Medicare Payment Policy.* Washington, DC: MedPAC.

———. 2003. *Report to the Congress: Medicare Payment Policy.* Washington, DC: MedPAC.

———. 2004. *Report to the Congress: Medicare Payment Policy***.** Washington, DC: MedPAC.

Mendelson, D., R. Ramchand, R. Abramson, and A. Tumlinson. 2002. *Prescription Drugs in Nursing Homes: Managing Costs And Quality in a Complex Environment*. National Health Policy Forum. NHPF-784. Washington, DC: The George Washington University.

Newhouse, J.P., S. Cretin, and C.J. Witsberger. 1989. "Predicting Hospital Accounting Costs." *Health Care Financing Review* 11(1): 25–33.

Office of Inspector General (OIG). 2001. *Medicare Beneficiary Access to Skilled Nursing Facilities*. OEI-02-01-00160. New York: Department of Health and Human Services.

Prospective Payment Assessment Commission (ProPAC). 1987. *Report to the Congress*. Washington, DC: ProPAC.

Rogowski, J.R., and D.J. Byrne. 1990. “Comparison of Alternative Weight Recalibration Methods for Diagnosis-Related Groups.” *Health Care Financing Review* 12(2): 87-101.

Schlenker R.E., A.M. Kramer, C.A. Hrinevich, and T.B. Eilertsen. 1997. “Rehabilitation Costs: Implications for Prospective Payment.” *Health Services Research* 32(5): 651–88.

Skilled Nursing Facility Prospective Payment System Technical Advisory Panel (SNF TAP). 2004. *Meeting (June 2004) and Summary of Survey Findings (August 2004).* Centers for Medicare and Medicaid Services Contract no. 500-00-0025/TO#2. Baltimore, MD: Centers for Medicare and Medicaid Services.

Urban Institute, The. 2004. *Classifying Medicare SNF Patients: Initial Findings from Three Approaches.* Report to the Centers for Medicare and Medicaid Services. Contract no. 500-00-0025. Washington, DC: The Urban Institute.

White, C., S. Pizer, and A. White. 2002. “Assessing the RUG-III Resident Classification System for Skilled Nursing Facilities.” *Health Care Financing Review* 24(2): 7–15.

Appendix to Chapter II

Appendix II.1: Variance Explanation of RUG-III

The predictive performance of RUG-III has been examined mainly in papers describing its development and refinement. In Fries and colleagues' 1994 description of developing RUG-III, the system explained 55.5% of the variance in nursing and therapy per diem resource use (measured as staff time), and 41.2% of the variance in nursing resource use alone.
After implementation of the SNF PPS, attention was focused on RUG-III's variance explanation regarding Medicare patients (rather than all nursing facility patients). In a subsequent analysis of the sample used in the development of RUG-III, other researchers reported that RUG-III explained 39% of variation in routine and therapy resource use (staff time minutes) among Medicaid residents and 21% of variation in staff time minutes among Medicare SNF patients (Abt Associates 2000). The Abt researchers posited that RUG-III explained less variation among Medicare than Medicaid staff time minutes in part because the majority of Medicare patients were classified in a single RUG-III category (rehabilitation), and the resulting case-mix heterogeneity and level of resource use variation in that category decreased the system's predictive capability regarding the Medicare population. About 65% of Medicare stays were in that category, in contrast to only 18% of Medicaid stays.

In another study of the RUG-III system, Kramer and colleagues (1999) collected patient assessment and staff time data in 1998 on a purposive sample of 1,000 Medicare patients with high rehabilitation and medically complex needs in 64 SNFs across 14 states. In this resource-intensive patient sample, the variance explanation of RUG-III was dramatically lower (8%, nursing and therapy staff time; 4%, nursing staff time; 21%, therapy staff time).
Our analyses of 1999 Medicare SNF claims also indicate low variance explanation (5% regarding total cost) when using Medicare costs (derived from charges), rather than staff time, as the measure of resource use (Liu, et al. 2002). When using this measure of resource use, the variance explanation regarding routine costs is low (3% in our analyses), because this measure captures variation in routine costs only at the facility, rather than patient, level. Variance explanation of NTA costs also is low (5% in our analyses). This finding is not surprising because RUG-III was not developed to predict NTA costs. As expected, the explanation is highest regarding therapy costs (38% in our analyses).

Appendix to Chapter III

Appendix III.1: Data and Methods

Data Sources

The MDS is part of an overall nursing home resident assessment system, required by the Nursing Home Reform Act of OBRA 1987, which was developed to improve the health and quality of life of nursing home residents. The MDS consists of over 300 questions grouped into 18 domains—such as patient diagnosis, cognitive status, functional status, and nutritional status. Of the roughly 300 items on the full assessment form, about 100 are used for PPS payment purposes. In DataPRO, all MDS assessments are matched to each SNF stay, with selected items pulled for the file. MDS assessments for SNF patients are made on a specified schedule about 5, 14, 30, 60, and 90 days from the start of the Medicare stay.

Medicare certified providers submit claims for reimbursement for Medicare-covered services, with institutional providers using form UB-92. Information on these claims includes periods of service, types of procedures furnished, primary and secondary diagnoses of the patient, and the institution's charges for services provided. Claims are submitted to Medicare intermediaries who conduct edits, review for appropriateness, and reimburse the provider according to Medicare coverage, eligibility, and payment rules. Claims from the Medicare intermediaries are ultimately sent to CMS.

Construction of the 2001 Analysis Files

We used the DataPRO SNF and prior hospital stays as the starting point for creating the analysis file. We examined the stays to identify potential integrity problems and non-Medicare coverage, and dropped cases with anomalies (such as overlapping claims records, zero covered days, and missing prior hospital stay information). We then examined the MDS assessments associated with each stay and dropped cases with anomalies (such as absence of any assessments, irregular patterns of assessments, and mismatches between MDS assessments in DataPRO and MDS assessments from the MDS core data files). Next, we examined consistency of information between DataPRO stays and the SNF stays we created from raw claims data (which would be the source of service use information not originally collected by DataPRO). Medicare stays found to have inconsistent information were excluded from the analysis files.

The next step was to determine the intersection between the "clean" Medicare SNF stays and the SNF cost reports on record. SNF stays without corresponding cost report data—which could not have the claims amounts converted to estimated costs—were also excluded from the analysis files. The effects of the exclusions on the number of cases are summarized in Appendix Exhibit III.1.

Appendix Exhibit III.1: The Number of Cases Remaining after Specific Initial Edits

Reasons for Exclusion of Stays	***Remaining Numbers***
Total SNF stays in 2001 DataPRO	2,114,797
1. "Integrity Problems and Non-Medicare coverage" • Overlapping SNF stays, fragmented SNF stays, overlapping claims records within stays, no Medicare payment, no covered days • Overlapping qualifying hospital stays, no qualifying hospital stay, overlapping claims records for the qualifying hospital stay, no Medicare payments • Non-PPS stay • Swing Bed stay	1,900,036
2. Unavailability of MDS assessments • No MDS assessments • No 5-day MDS assessment • Missing complete sequence of assessments (e.g., a stay with 5-day and 30-day assessments, but no 14-day assessment) • Mismatch between SNF stays' assessments and base MDS data set	1,778,059
3. Mismatch of the DataPRO stays and supplemental claims data • Non-match of stays (SNF stays or hospital stays) • Stays with different SNF covered days • Stay with different hospital covered days	1,768,761
4. Matching with cost report information and wage-index file	1,722,987
5. Other exclusions • Length of stay is not equal to Medicare covered days • Stays that have the same assigned MDS assessments	1,709,736

Facility Data Exclusion Rules (Flags)

We identified facilities whose data were likely to be too unreliable to be included in an analysis sample. We encountered three types of problems on facility cost reports that led to the creation of flags for specific facilities: (1) unreliable cost data on the cost report; (2) an inability to accurately assign Medicare costs from the total facility costs; and (3) unreliable cost to charge ratios.

Unreliable Cost Data

Cost data were determined to be unreliable based upon unlikely cost breakdowns or extreme cost levels. We flagged facilities that had unacceptable ratios of ancillary to routine costs based upon ratios less than or equal to .005 or ratios greater than 5.0. Inconsistent facilities with total ancillary costs that exceeded the sum of the costs of the ancillary cost components by a significant amount were identified as unreliable. We also flagged four facilities with the following extreme costs: routine cost per day of $0.009, routine cost per day of $36,409, routine cost per day of $0.28, and non-therapy ancillary cost per day of $8,815.

Inability to Assign Medicare Costs in a Facility

We needed to assign Medicare costs in a facility in order to generate cost to charge ratios for calculating costs for the observations in the file. Facilities that were missing data necessary to assign Medicare costs were flagged. These included facilities missing any of the following data: number of SNF beds, number of Medicare patient days, number of SNF participating unit days, and number of days in other units of the nursing home.

Unreliable Cost to Charge Ratios

We conducted an analysis to determine which cost to charge ratios (CCRs) to categorize as “out of range.” We flagged SNFs with such ratios for any of the service components we examined. It is important to note that there is no “right” answer; this analysis simply served to help us make decision rules. Our goal was to be as inclusive as possible, while identifying particular SNFs that were likely to contribute clearly erroneous data to our classification analysis.

We first explored the ranges of CCRs used by other analysts in Medicare PPS-related research. In research on inpatient rehabilitation facilities, RAND settled on a range of 0.05-10.0 for departmental CCRs; when values were outside this range, values were imputed based on like facilities. In earlier research on SNF refinement, Abt Associates used a range of 0.25-2.1 for ancillary services. In research on acute care hospital CCRs, Newhouse, et al. (1989) used limits for departmental CCRs that varied by size of hospitals: 0.01- 100 for hospitals with less than 100 beds and 0.01-3.0 for hospitals with 100 or more beds.

We conducted detailed analyses of CCR distributions separately for hospital-based and freestanding SNFs, flagging SNFs with CCRs that appeared to be extreme values. From this analysis, we determined that a range 0.05 – 30.0 was a reasonable starting point for (a) total ancillary, (b) rehabilitation therapy, (c) total NTA, and (d) drugs. For respiratory therapy and other NTA services, we determined that a higher cap would be appropriate. We selected the range 0.05 to 100 for these last two components for freestanding SNFs. We flagged hospital-based SNFs with service specific CCRs outside a 0.05-10.0 range for all components.
The impact of applying these flag rules on number of SNFs and number of associated stays, separately for freestanding and hospital-based SNFs, is relatively minor. For freestanding SNFs, the 112 flagged SNFs and 1,175 associated stays are 0.9% of all freestanding SNFs and 0.9% of all stays in freestanding SNFs. For hospital-based SNFs, the flagged cases are 0.8% of all hospital-based SNFs and 0.3% of all stays in hospital-based SNFs.

We determined that the "generous" range of CCRs we applied included virtually all stays in freestanding and in hospital-based SNFs. This high yield led us to exclude the remaining small number of stays from the analysis without attempting to impute values for them.

A related issue is that some claims are submitted for particular types of services (e.g., respiratory therapy), for which CCRs could not be constructed using a facility's cost reports. In these cases, we applied the available CCR from the next higher level of service aggregation. For example, if a claim for respiratory therapy was submitted but the SNF did not have a respiratory CCR, we applied the CCR for total NTA. Similarly, if a claim for rehabilitation therapy was submitted but no rehabilitation therapy CCR existed, we applied the CCR for total ancillary services.

Stay Data Exclusion Rules (Flags)

Finally, we eliminated stays with extreme values for total ancillary costs and charges out of concern for the validity of the data. We dropped stays with logged total ancillary costs or charges 3 standard deviations from their logged mean, similar to previous work on hospital costs. Since most of the stays identified through costs overlapped those identified with extreme charges, the process eliminated only about 2.0% of the remaining observations.

Selecting Units of Observation

The key limitation in the analysis file data is that ancillary service charges, and, thus, costs, are not reported for specific dates of service, but have to be averaged and applied to the entire SNF stay. Hence, the MDS assessments on patient conditions, which are recorded at specific dates during a stay, cannot be linked directly to specific cost information. Because about 60% of the SNF stays have multiple MDS assessments, it is important to consider if, and how, information from all the assessments can be used in the prediction of costs.

We developed a procedure to make efficient use of all the MDS assessments for each stay by defining segments for each MDS and weighting them together based on the proportion of time during the stay covered by each assessment. The reasoning behind this approach is that the outcome – per diem cost measured over the entire stay – is a weighted average of the per diem costs measured over the MDS assessment periods, or segments. As a result, it is appropriate to examine cost variation with the weighted average of the patient conditions over the same assessment periods, or segments.

Appendix III.2: Statistical Evaluation of Patient Classification Models

This appendix describes the general features of the models used in each patient classification approach and the minimum common set of statistics that are reported and used to evaluate each approach.

Patient Stay-Level Models and Statistics

Classification systems used in a PPS should be able to account for a reasonably high proportion of the predictable variation in a provider's patient care costs due to clinically meaningful differences in patient characteristics. To the extent that a classification system does not sufficiently meet this goal, provider incentives to select patients according to risk may increase. Predicting variation in costs that is due to clinically inappropriate variation (e.g., provider inefficiencies or regional practice patterns not related to best practices) is not desirable. In other words, higher predictability of cost variation is generally considered better, as long as it is the "right kind" of variation. The unit of payment currently used for SNFs under Medicare is a day of patient care. Thus, a classification system should explain sufficient appropriate variation in costs per day at the person level.

Patient stay-level models are estimated using both costs per day and charges per day as dependent variables. Costs and charges represent two alternative measures of SNF resource use. Because both measures have conceptual advantages and disadvantages, we consider the 2 measures separately to allow comparisons of the results obtained under each measure using a particular classification system. The cost and charge variables we use in the stay-level models have been wage-adjusted using the area Wage Index Factor (WIF) and a labor share of 0.7787, unless stated otherwise. For simplicity, we focus our discussion here primarily on costs, even though analogous models are examined using charges.

The general form of the stay-level model is:

$$f(\textit{wage-adjusted cost per day}_i) = g(\textit{patient classification system variables}_i),$$

where f and g are functions and i denotes a patient stay. We leave the functional form arbitrary because models may be estimated in a variety of ways, including linear regression on costs, linear regression on the log of costs, two-step regression, and categorical regression tree analysis (CART).

R-squared

The R-squared statistic reported for each model tells us what share of the variation in the dependent variable (e.g., costs per day) can be accounted for by the patient classification system. In linear regression models, the usual R-squared is reported (explained sum of squares / total sum of squares). For two-step or uncentered models (where the average predicted cost is not equal to the average actual cost) and for out-of-sample predictions, we report the R-squared as computed by a regression of the actual cost on the predicted cost. Regression results are also reported if the coefficient on the prediction differs significantly from 1.0 or if the coefficient on the constant term differs significantly from 0. The constant term of this regression measures the degree of uncenteredness.

In comparing models with the same dependent variable but different sets of explanatory variables, so long as the cost variation being explained is actually attributable to clinically appropriate patient characteristics, the classification system with the higher R-squared is preferred, all else equal. It is important to recall that patient characteristics may be correlated with facility characteristics that are not appropriate for payment adjustment, however, which may contribute to a higher R-squared for unintended reasons. Another point to keep in mind is that the R-squared cannot be compared unambiguously across models estimated using different samples, different dependent variables, or different transformations of dependent variables. For example, a higher R-squared using charges as the dependent variable, rather than costs, does not necessarily imply that the charge model is a better model of patient resource utilization. R-squared is an overall, or summary, measure of predictive ability. Two models with similar R-squared could differ in how well they predict cost for high cost patients versus low cost patients. We can obtain more detail from measuring sensitivity.

Sensitivity

Sensitivity measures how well the classification system can correctly predict high cost cases to be high cost. We define high cost cases for this purpose as cases in the highest 10% of costs. Sensitivity is calculated as the percent of cases in the highest 10% of actual costs that are also in the highest 10% of predicted costs. As such, the sensitivity measures tell us the probability that the most expensive stays will be paid at the highest payment rates under a given classification system. All else equal, a more sensitive classification system will be less likely to create incentives for providers to select against the most costly patients.

Standard Deviation of the Relative Weights

A patient classification system, implemented as a set of categories or a regression-based index model, yields a set of predicted costs that depend on patient characteristics. Relative weights (also called relative values) measure the costliness of a patient type relative to the average patient. Relative weights are constructed by dividing the predicted cost for a particular type of patient by the average actual cost (or average predicted cost).[29] The relative weights will have a mean of 1.0 over the whole estimation sample either by design or, if necessary, by re-centering them. Relative weights will normally have an average close to, but not exactly, 1.0 when computed for randomly selected sets of patients outside the estimation sample.

For each of the base case models, we report the standard deviation of the relative weights, which provides an indication of the extent of variability of the payment rates that will result from a particular payment system. There is no "correct" value for this measure, but it provides a gauge of how different the dispersion in payment rates would be under different patient classification systems. Our primary purpose in reporting this measure is to understand the extent to which variation in payments would differ for a given classification system depending on whether costs or charges are used as the dependent variable.

Validation Out-of-Sample

The base case models are estimated using the test sample (i.e., the 10% simple random sample of stays). The statistics listed above are reported for each model. For out-of-sample validation, the results (e.g., regression coefficients) of these models are then applied to a new sample of stays: the validation sample (i.e., the stays that constitute the 10% facility-stay file). This process mimics the inherently out-of-sample use of classification systems in practice, whereby payment rates developed from older data are used for determining current payments.

We report each of the stay-level model statistics separately for stays in the validation sample. It is reasonable to expect minor deterioration of model results when applied to the validation sample. Large differences between estimation sample and validation sample results, however, may indicate substantial over-fitting of models to the data in the estimation sample. Because the statistics presented for the validation sample provide a more realistic appraisal of the statistical properties of the classification system, we rely on the validation sample results in comparing the different approaches.

Facility-Level Models

Effective classification systems should also be able to account for variation in costs across facilities that are attributable to variation in average patient case-mix characteristics. Facilities may have high average costs due to a more severe or complex patient mix, perhaps because they are well equipped to treat such patients. Facilities with a more severe case-mix would be penalized if the patient classification system does not reflect these differences such that financial losses on high cost patients sufficiently offset gains from low cost patients (MedPAC 1999). Differences in facility payment levels due to case-mix differences should both explain and be proportional to facility differences in average costs per patient day in SNFs.

---

[29] The average actual cost and average predicted cost are the same in one-stage regression models that contain a constant term.

We estimate facility-level cost models to address these issues, examining how facility costs vary with case-mix-adjusted payments and other facility characteristics. We use the validation sample for this purpose. Stay-level costs are averaged, weighting by Medicare covered patient days, to construct the average cost per patient day in each facility.

Facility Case-Mix Indices

Facility case-mix indices (CMIs) summarize the costliness of a facility's patient case-mix as captured by a particular classification system. CMIs are constructed by averaging the (out-of-sample predicted) relative weights of the stays within each facility in the validation sample. The average CMIs are weighted by the number of covered days for each stay. At the facility level, CMIs are re-centered so that the means of the CMIs across all facilities are equal to 1.0. We estimate facility-level models by regressing the log of facility average costs per day (unadjusted for area wages) on a constant, the log of the CMI, and, possibly, on a set of control variables, which varies according to the model specification. We estimate and report 3 specifications using the base case from each classification approach.

CMI-only model:

$$\log(\textit{Facility cost per day}) = \beta_0 + \beta_1 * \log(\textit{CMI}) + \varepsilon_1$$

Payment model:

$$\log(\textit{Facility cost per day}) = \gamma_0 + \gamma_1 * \log(\textit{CMI}) + \gamma_2 * \log(\textit{WIF}) + \gamma_3 * \textit{Urban} + \varepsilon_2$$

Fully-specified model:

$$\log(\textit{Facility cost per day}) = \delta_0 + \delta_1 * \log(\textit{CMI}) + \delta_2 * \log(\textit{WIF}) + \delta_3 * \textit{Urban} + X\delta_4 + \varepsilon_3.$$

The CMI-only model contains no control variables except for a constant term. The payment model adds the log of the area wage index factor (WIF) and an urban facility indicator as control variables, because these 2 variables are currently accepted as valid facility characteristics for which SNF payments should be adjusted. The "fully-specified" model adds control variables, denoted by *X*, for number of beds (0 to 9 beds, 10 to 19 beds, 20 to 49 beds, 50 to 99 beds, with 100+ beds as the excluded category); being a hospital-based facility; ownership type (voluntary/nonprofit, for profit-other, and government, with for profit-chain as the excluded category); and percent of SNF patient days paid by Medicare. These additional variables are expected to be correlated with facility costs and may be correlated with the CMIs and unobserved aspects of patient case-mix; however, they are characteristics that are not considered appropriate to use in adjusting SNF costs *per se*.

The CMI variables may be based on cost- or charge-based relative weights. Regardless of which is used, the dependent variables are facility costs (not charges) because costs are conceptually the more accurate measure of SNF resource use. The potential problem with estimated costs at the stay level, which motivates our consideration of charges (that cost-to-charge ratios must be assumed fixed at the facility level), is not an issue at the facility level, because SNF costs are directly measured at the facility level from the Medicare cost reports. In evaluating the facility-level models, the main statistics of interest, reported for all base models, are the model R-squared and the coefficients on the logs of the CMIs in each specification.

Facility-Level Model R-squared

The R-squared of the CMI-only model indicates the maximum share of variation in (log) average SNF costs per day that can be attributed to facility case-mix as measured by the particular classification system. This measure can be compared across different classification systems for the same cost component. It is also informative to examine how the R-squared value changes as additional explanatory variables are added. The change in R-squared from the CMI-only model to the payment model summarizes the potential contributions of further payment adjustments for area wage levels and urban status. The change in R-squared from the payment model to the fully-specified model summarizes the unique contribution of the other facility characteristics that are considered inappropriate for payment adjustment in explaining SNF costs. The larger the contribution of these non-payment variables is relative to the R-squared in the payment model, the larger the potential for important financial impacts across facilities of different types. Even financial impacts that may be considered desirable in the long run if they strengthen the overall efficiency of service provision may be disruptive to patient care in the short run if they lead high cost facilities to close.

The CMI Coefficient and CMI Compression

With the log-log model specification, the CMI coefficient (i.e., the regression coefficient on the log (CMI) variable in a facility-level model) measures the percent change in expected facility costs associated with a percent change in the CMI (and, thereby, its average relative payment rate per patient day due to case-mix). The ideal situation is for the CMI coefficient to be 1.0, indicating, for example, that a 10% higher CMI is associated with 10% higher costs. When the CMI is proportional to expected costs, facilities are paid in proportion to the higher cost associated with a higher CMI.

The situation that arises with a CMI coefficient that is greater than 1.0 is called “CMI compression.” If, for example, the estimated CMI coefficient is 1.25, a facility whose CMI and average payments are 10% higher than average would tend to have costs that are 12.5% higher than average. Costs tend to rise (and fall) more than payments, so that high case-mix facilities tend to be underpaid and low case-mix facilities tend to be overpaid. The term compression is used because the distribution of case-mix-adjusted payments is more narrow, or compressed, than the distribution of expected costs associated with facility case-mix. A classification system that exhibits CMI compression would tend to penalize facilities that have a more severe case-mix and create incentives to avoid high cost patients (Cotterill 1986).

A CMI coefficient that is less than 1.0 would indicate what is called "CMI decompression."[30] Decompression is a less common finding in the literature, but it arises in some of the results we present below. Decompression appears more prevalent in our analyses that use charges as the measure of resource use. A possible explanation (requiring further investigation) is that charge-based relative weights may be partly picking up charging practices, which get incorporated into the CMIs and lead to reduced CMI coefficients once facility characteristics are added as controls. When decompression is present, facilities with a CMI greater than 1.0 will tend to be overpaid and facilities with a CMI less than 1.0 will tend to be underpaid.

We illustrate the payment implications of combinations of CMI compression, proportionality, and decompression for high, average, and low CMI facilities in Appendix Exhibit III.2. For this illustration, we use a CMI-only model in which the average facility CMI is 1.0 and the average cost per day is $100. To take just the decompression example, in which we show the implication of having a CMI coefficient of 0.8, a facility with a CMI of 1.4 would make an average profit of $9.11 for each patient day. A facility with a CMI of 1.0 would break even. A facility with a CMI of 0.7 would lose $5.18 per patient day.

[30] The term "decompression" could also be used to describe a procedure to adjust for compressed weights, but here we mean the opposite of compression.

Appendix Exhibit III.2: Illustration of CMI Coefficient and Facility CMI Scenarios

Situation	CMI coefficient (β_1)	Facility CMI	Expected log facility cost per day	Expected facility cost per patient day	Facility payment per patient day	Profit (loss) per patient day
Compression						
High case-mix facility	1.25	1.4	5.03	$152.29	$140	$(12.68)
Average case-mix facility	1.25	1.0	4.61	$100.00	$100	$0.00
Low case-mix facility	1.25	0.7	4.16	$64.03	$70	$6.47
Proportional						
High case-mix facility	1.0	1.4	4.94	$140.00	$140	$0.00
Average case-mix facility	1.0	1.0	4.61	$100.00	$100	$0.00
Low case-mix facility	1.0	0.7	4.25	$70.00	$70	$0.00
Decompression						
High case-mix facility	0.8	1.4	4.87	$130.89	$140	$11.68
Average case-mix facility	0.8	1.0	4.61	$100.00	$100	$0.00
Low case-mix facility	0.8	0.7	4.32	$75.18	$70	$(7.44)

Example assumptions: Log(average cost per day) = β_0 + β_1*Log(CMI). Average CMI = 1.0. Average cost per day = $100. β_0 = Log(100) = 4.605.

We report the CMI coefficient for each of the 3 specifications. In addition to the CMI, we report the standard error of the CMI coefficient, so that confidence intervals can be computed, and the p-value for the t-test of the null hypothesis that the CMI coefficient is equal to 1.0. The CMI coefficient results for the CMI-only specification are the least relevant when we consider that payment adjustments are made for area wages and urban status. They are reported primarily for completeness, although they are also useful to illustrate whether wage and urban adjustments alter the relationship between the CMIs and costs.

Interpreting Compression Results for the Payment and Fully-Specified Models

Findings of CMI compression or decompression have somewhat different meanings and implications in the payment and fully-specified models. The CMI coefficients in the payment models quantify the relationship between facility case-mix and costs across facilities, holding area wages and urban status fixed (or as if an average cost adjustment has already been made for those 2 factors). The CMI coefficient tells us whether case-mix adjusted payments will flow to facilities proportionately to their expected costs.

The fully-specified models estimate the relationship between case-mix-adjusted payments and costs within classes of facilities defined by the additional control variables. As such, the CMI coefficients measure the extent to which facilities of a given type would face incentives to risk-select patients on the basis of the characteristics included in the classification system. Facilities of a given type may be making or losing money overall. However, if costs are related proportionately to the CMI in the fully-specified models, they will not be making or losing more or less money depending on the case-mix characteristics of their patients.
If the fully-specified model is in fact a reasonably accurate overall model of the drivers of SNF costs, compression in the fully-specified models would cast serious doubt on the effectiveness of a classification system that is not otherwise modified with payment cushions, such as outlier payments. Decompression would also suggest serious problems that an outlier policy would not address.

Appendix Exhibit III.3 shows the 9 possible scenarios for compression, proportionality, and decompression in the payment and fully-specified models and their basic interpretation. Compression in the fully-specified model is very likely to be accompanied by compression in the payment model (since CMI coefficients usually fall when additional control variables are added to the model). It would, therefore, be unlikely that we would find a higher CMI coefficient in the fully-specified model than in the payment model; thus, we focus on the more likely scenarios.

Appendix Exhibit III.3: CMI Coefficient Scenarios in Payment and Fully-Specified Facility Models and Implications

	Fully-specified model		
Payment model	CMI coefficient > 1 Compression	CMI coefficient = 1 Proportionality	CMI coefficient < 1 Decompression
CMI coefficient > 1 Compression	Problematic • High cost facilities' costs exceed payments and low cost facilities' payments exceed costs • Incentive to select against high cost patients	Ambiguous • Depends on reason for compression • High cost facilities' costs exceed payments and low cost facilities' payments exceed costs • No incentive to select patients based on characteristics in classification system	Problematic • Indication of powerful facility effects • High cost facilities' costs exceed payments and low cost facilities' payments exceed costs • Incentive to select against low cost patients
CMI coefficient = 1 Proportionality	Unlikely	Ideal • Facilities' case-mix-adjusted payments equal to expected costs • No incentive to select patients based on characteristics in classification system	Problematic • Facilities' case-mix-adjusted payments equal to expected costs, but • Incentive to select against low cost patients
CMI coefficient < 1 Decompression	Unlikely	Unlikely	Problematic • High cost facilities' payments exceed costs and low cost facilities' costs exceed payments • Incentive to select against low cost patients

Proportionality in both equations is ideal and would suggest both payment equity across equations and minimized incentives to risk-select on the basis of factors included in the patient classification system.

One likely case with an ambiguous interpretation is a finding of proportionality in the fully-specified model but compression in the payment model. Under this scenario, the classification system may or may not be adequate, depending on the most likely source or sources of the compression. A specific example best illustrates this case. Suppose we find a CMI coefficient of 1.2 in the payment regression and 1.0 in the fully-specified regression. This result could only occur if the added facility characteristics in the fully-specified model were related both to costs and to the CMI. Assume the CMIs are related to facility characteristics simply because different types of patients are more appropriately treated in different types of facilities. In this case, interpreting the differences in CMI coefficients requires an understanding of, or assumptions about, why the facility characteristics relate to costs. This can be examined statistically.

Facility characteristics could be related to costs for some combination of 4 main reasons:

- Cost reporting or accounting differences across facility types
- Case-mix differences across facility types not captured by the classification system
- Operational efficiency differences by facility type
- Quality of care differences by facility type

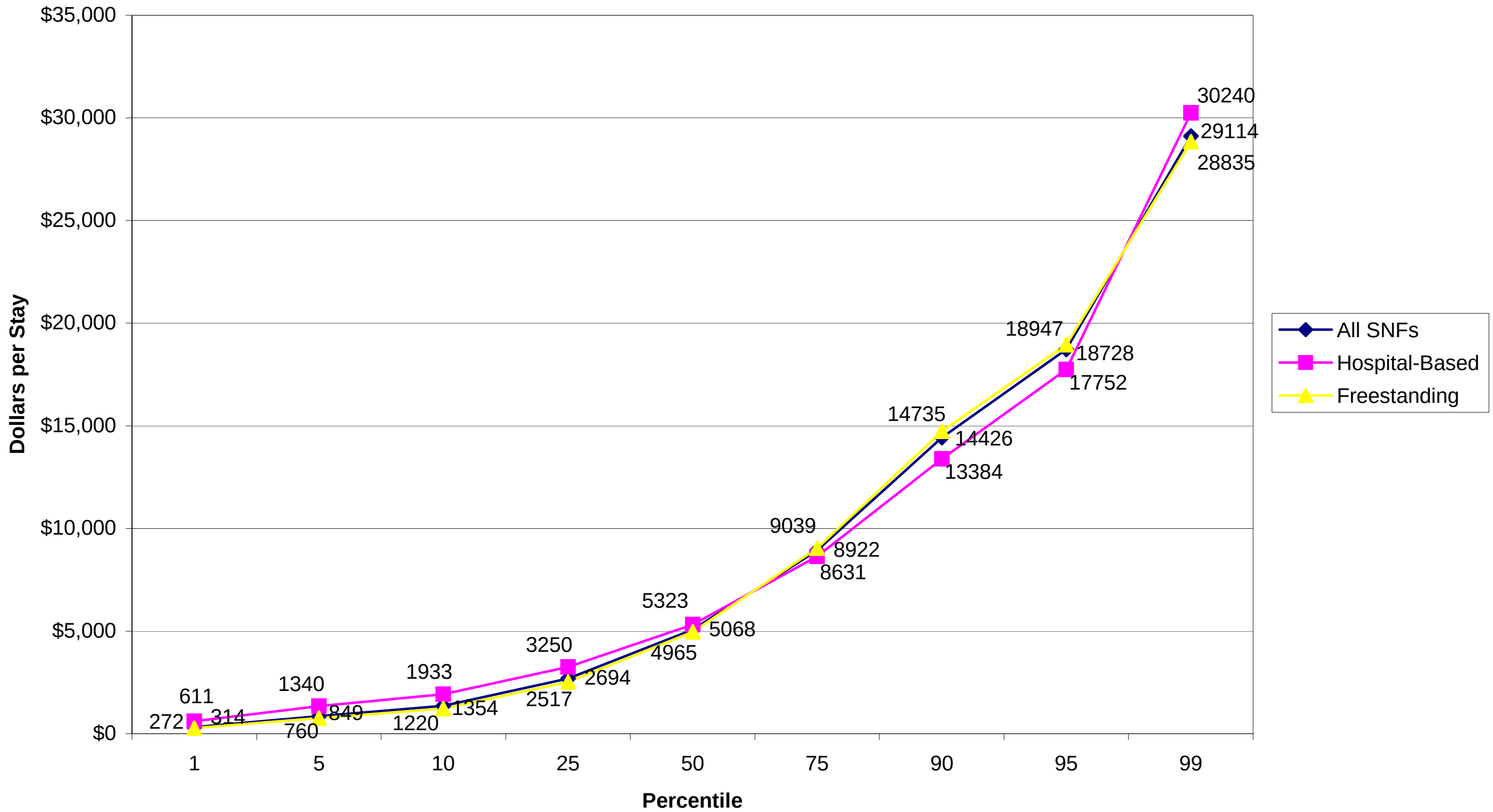
Each reason would have a different implication for the compression in the payment regression, and, although we lack sufficient information to know the *relative* contribution of each with any degree of certainty, we can at least consider their separate implications.
If compression in the payment model were driven by cost accounting differences across facility types, there would be no concern. Facilities would receive case-mix adjusted payments proportionate to their true underlying costs.

If the compression resulted from case-mix differences arising from type of facility, but those differences not captured by the classification system and thus not reflected in payments, it would be problematic. Costly case-mix facilities would not be paid sufficiently or fairly for their more costly case-mix and low cost facilities would be overpaid.
If the compression was due to inefficiency in certain types of facilities, inefficient facilities would have an incentive to become more efficient, which would further the goals of the Medicare program in the long run. However, financial strain on inefficient providers could potentially disrupt patient care in the short run, should providers close or reduce their level of patient care.

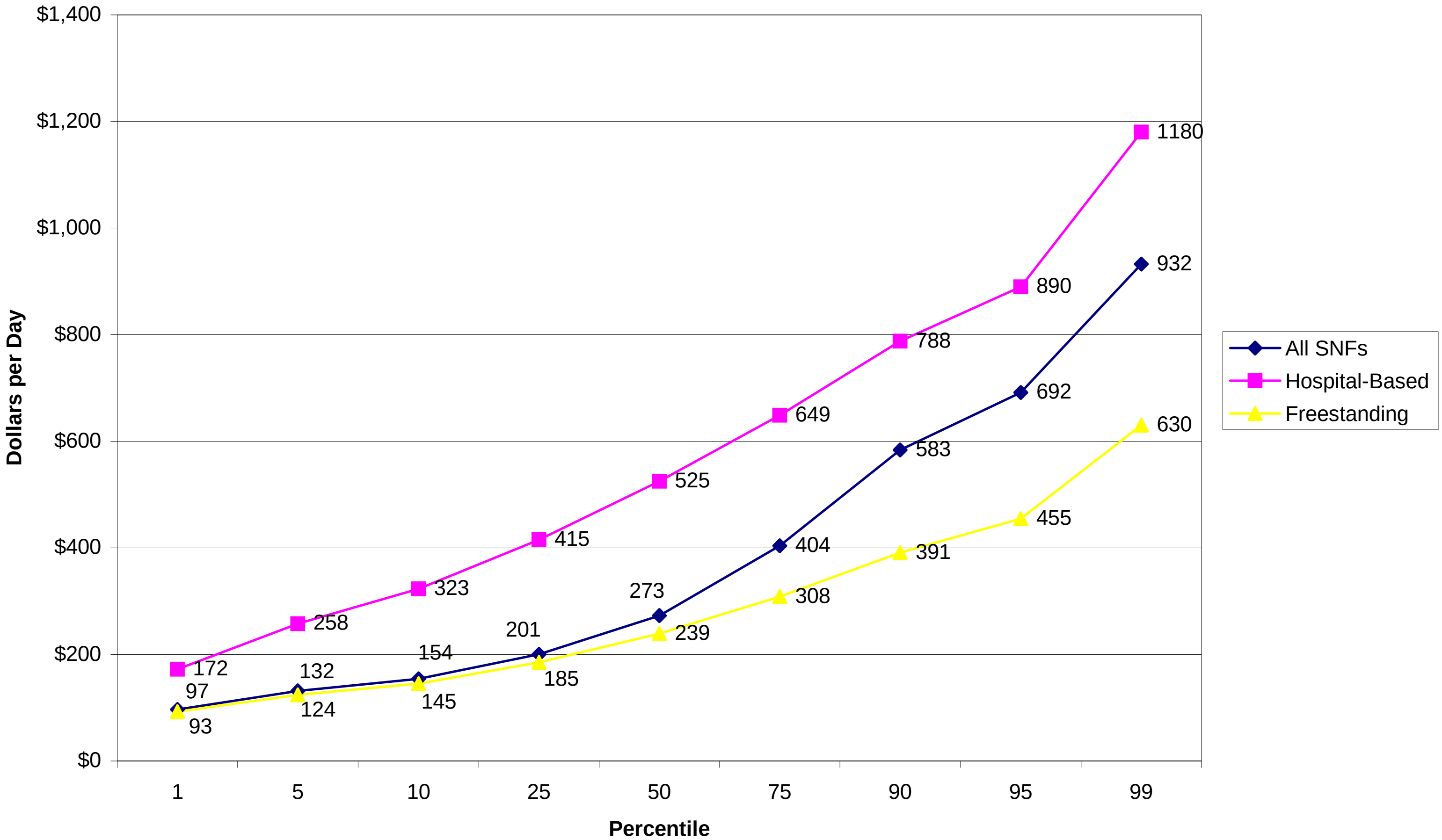
If the compression resulted from differences by facility type in the quality of SNF patient care, while quality-adjusted costs would vary in proportion to payments, high quality facilities would receive no reward for superior quality and low quality facilities would incur no penalty for providing inferior care. Policymakers would need to balance patient care objectives against fiscal objectives, and there is currently no provision for compensating higher quality SNF care through Medicare payment adjustments.

Appendix to Chapter VIII

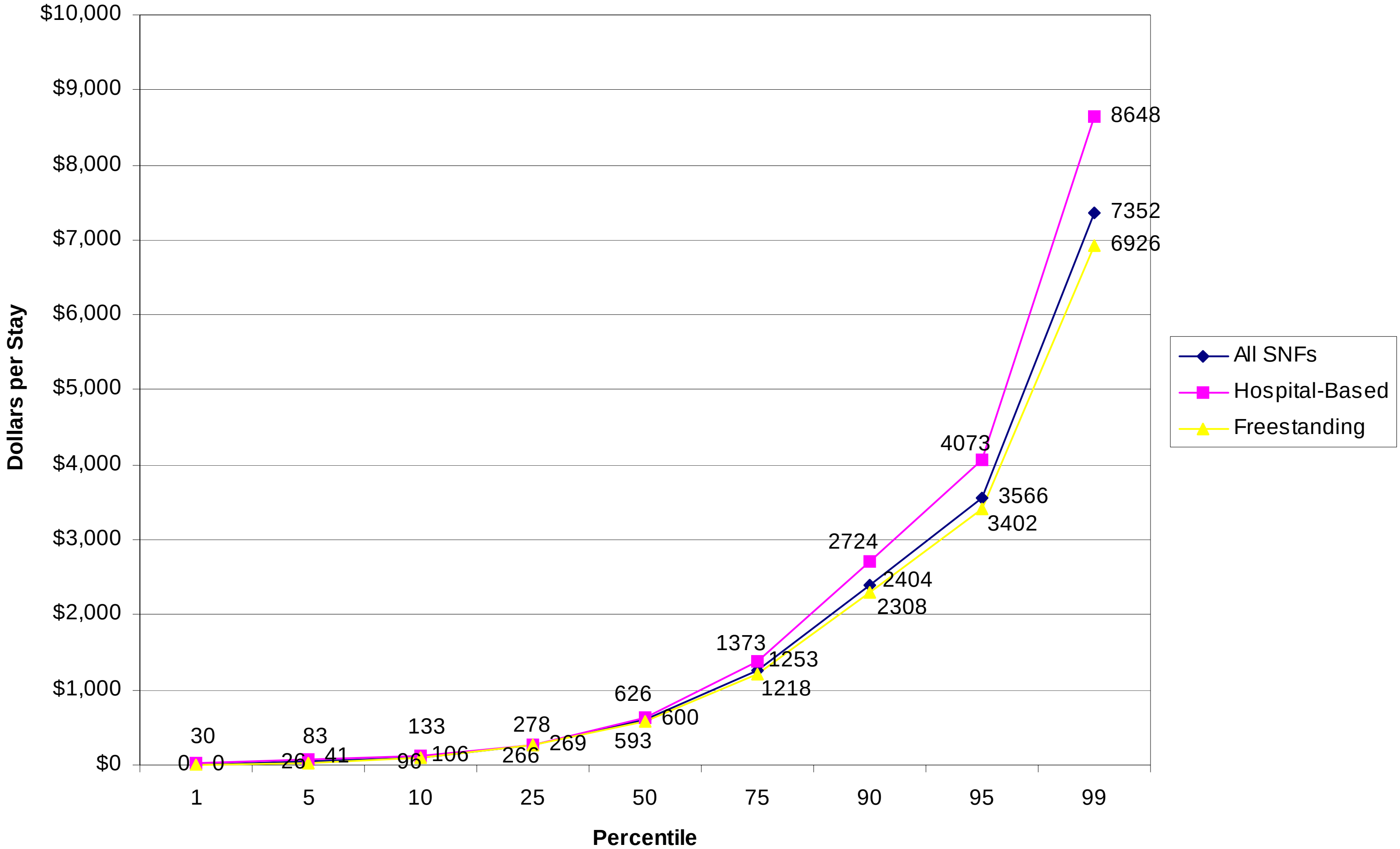
Appendix Exhibit VIII.1: Distribution of SNF Total Costs per Stay, 2001



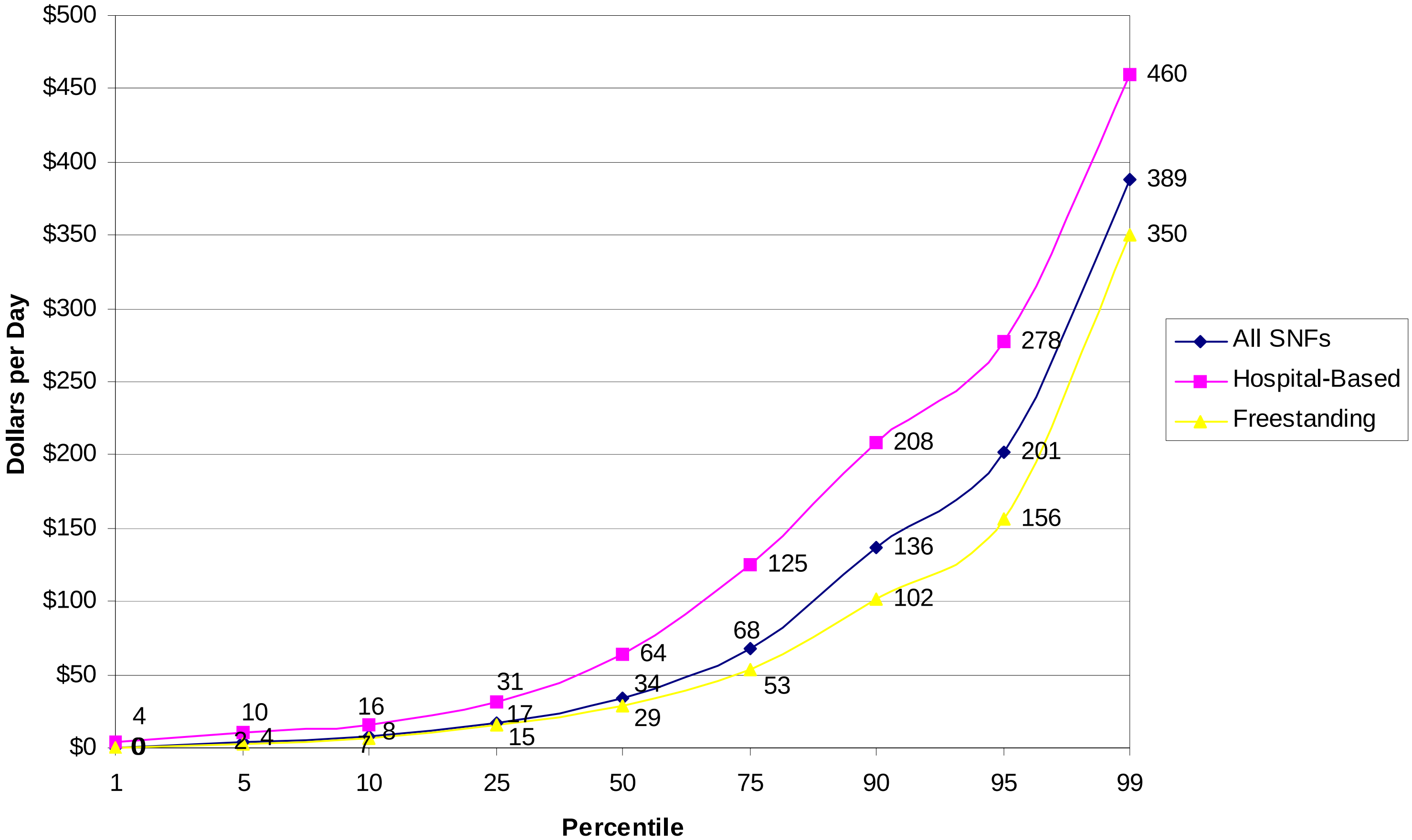
Appendix Exhibit VIII.2: Distribution of SNF Total Costs Per Day, 2001



Appendix Exhibit VIII.3: Distribution of SNF NTA Costs Per Stay, 2001



Appendix Exhibit VIII.4: Distribution of SNF NTA Costs per Day, 2001



Appendix Exhibit VIII.5: List of States by Census Division

Census Division	State	Census Division	State
New England	Maine	West North Central	Minnesota
	Vermont		Iowa
	Massachusetts		Missouri
	New Hampshire		Kansas
	Connecticut		Nebraska
	Rhode Island		North Dakota
			South Dakota
Mid Atlantic	New York		
	New Jersey	West South Central	Arkansas
	Pennsylvania		Louisiana
			Texas
South Atlantic	Delaware		Oklahoma
	Maryland		
	District of Columbia	Mountain	Montana
	Virginia		Wyoming
	West Virginia		Colorado
	North Carolina		New Mexico
	South Carolina		Arizona
	Georgia		Utah
	Florida		Idaho
			Nevada
East North Central	Ohio		
	Michigan	Pacific	Washington
	Indiana		Oregon
	Illinois		California
	Wisconsin		Alaska
			Hawaii
East South Central	Kentucky		
	Tennessee		
	Mississippi		
	Alabama		

Appendix Exhibit VIII.6: Outlier Stay and Payment Percentages under Five Total Outlier Payment Policies, by RUG-III Group

RUG-III Group	Outlier Stay Percents					Outlier Payment Percents				
	1%	2%	3%	4%	5%	1%	2%	3%	4%	5%
RUC	1.5	2.9	4.8	6.6	9.3	0.5	1.0	1.5	1.9	2.3
RUB	0.9	2.0	3.7	5.7	8.1	0.4	0.7	1.1	1.5	2.0
RUA	0.8	1.4	2.2	3.5	5.3	0.6	1.0	1.4	1.7	2.1
RVC	1.1	2.8	5.8	9.4	12.3	0.6	1.3	2.0	2.9	3.7
RVB	0.6	2.0	4.0	6.6	9.5	0.2	0.6	1.1	1.7	2.4
RVA	0.4	1.7	3.6	5.8	8.4	0.3	0.7	1.3	1.9	2.7
RHC	0.7	2.1	4.6	7.3	10.4	0.4	1.1	1.8	2.7	3.6
RHB	0.4	1.7	4.2	7.7	12.1	0.4	0.9	1.7	2.7	3.9
RHA	0.3	1.5	3.3	5.7	8.7	0.4	1.0	1.8	2.8	4.0
RMC	1.3	2.9	4.7	6.9	9.7	1.3	2.3	3.3	4.2	5.1
RMB	0.8	2.8	5.4	8.9	12.7	0.7	1.6	2.6	3.8	5.1
RMA	0.5	1.7	3.8	6.6	9.2	0.6	1.3	2.4	3.6	4.9
RLB	4.8	10.2	17.0	20.4	24.5	3.1	8.0	12.1	15.0	17.5
RLA	0.6	3.2	8.4	14.5	21.3	0.3	1.9	4.3	7.6	10.8
SE3	1.9	4.1	6.2	8.7	10.7	2.1	4.3	6.1	7.7	9.2
SE2	1.0	2.4	4.0	5.8	7.6	1.6	3.0	4.3	5.5	6.7
SE1	0.2	0.5	0.5	1.6	2.5	1.6	2.0	2.3	2.7	3.2
SSC	0.5	1.4	2.7	3.9	5.4	0.5	0.8	1.2	1.6	2.0
SSB	0.7	2.1	3.5	5.4	7.5	0.3	0.9	1.5	2.1	2.7
SSA	2.4	5.2	8.4	11.6	15.2	3.5	6.6	9.2	11.6	13.7
CC2	1.1	1.1	1.1	1.1	1.1	0.2	0.9	1.2	1.4	1.6
CC1	1.0	1.4	2.9	4.3	7.0	1.7	2.4	3.0	3.7	4.4
CB2	1.0	1.7	3.1	4.1	5.8	0.4	1.2	1.9	2.5	3.1
CB1	0.5	2.1	3.7	4.6	5.9	1.0	1.8	2.6	3.3	4.0
CA2	0.6	2.1	4.6	4.9	7.6	0.8	1.5	2.4	3.3	4.2
CA1	1.0	3.1	4.6	6.2	9.1	1.2	2.9	4.3	5.6	6.9
IB2	-	-	-	-	-	-	-	-	-	-
IB1	-	-	0.4	0.4	1.6	-	-	0.1	0.3	0.5
IA2	-	-	-	-	33.3	-	-	-	-	1.4
IA1	-	-	0.9	1.4	3.2	-	-	0.3	0.9	1.5
BB2	-	-	-	-	-	-	-	-	-	-
BB1	-	-	-	4.2	8.3	-	-	-	1.3	2.8
BA2	-	-	-	-	-	-	-	-	-	-
BA1	-	-	-	1.4	2.9	-	-	-	0.7	1.6
PE2	-	-	-	-	-	-	-	-	-	-
PE1	-	0.6	1.2	1.2	2.3	-	0.2	0.9	1.3	1.8
PD2	-	5.9	5.9	17.6	23.5	-	0.1	1.8	4.6	7.1
PD1	0.7	2.2	2.9	3.8	5.0	0.5	2.3	3.8	5.0	6.3
PC2	-	-	-	-	-	-	-	-	-	-
PC1	-	-	-	-	1.0	-	-	-	-	0.3
PB2	-	-	-	-	-	-	-	-	-	-
PB1	-	0.7	0.7	2.0	2.7	-	0.9	1.6	2.7	3.7
PA2	-	-	-	-	-	-	-	-	-	-
PA1	0.289	2.0	3.2	-	5.2	0.8	2.22	4.33	6.09	7.8

Note: Based on RUG-III assignments that reflect 75 percent or more of a stay. This restriction was used to allow outlier results associated with a RUG-III group to reflect stays that are generally characterized by that RUG-III assignment. Results for eight RUG-III groups are based on extremely small (n<30) sample sizes: IB2, IA2, BB1, PE2, PD2, PC2, PB2, PA2.

Appendix Exhibit VIII.7: Outlier Stay and Payment Percentages under Eight NTA Outlier Payment Policies, by RUG-III Category

RUG-III Category	Outlier Stay Percents								Outlier Payment Percents							
	1%	2%	3%	4%	5%	10%	15%	16%	1%	2%	3%	4%	5%	10%	15%	16%
Rehabilitation	0.1	0.3	0.5	0.8	1.2	3.9	7.9	8.5	0.3	0.8	1.3	1.9	2.5	6.4	11.1	11.9
Extensive Services	0.4	1.0	1.6	2.2	3.0	7.2	12.5	13.3	2.1	4.3	6.4	8.4	10.4	20.6	30.9	32.5
Special Care	0.4	1.5	2.3	3.4	4.3	10.4	17.6	18.6	3.6	7.1	10.7	14.1	17.5	33.9	50.0	52.4
Clinically Complex	0.2	0.6	1.0	1.4	1.9	5.0	10.1	10.9	1.2	3.3	5.2	7.0	8.8	18.2	28.4	30.0
Impaired Cognition	0.0	0.0	0.0	0.0	0.0	1.4	4.7	5.5	0.0	0.0	0.0	0.0	0.0	4.6	13.1	15.3
Behavioral Problems	0.0	0.0	0.0	0.0	0.0	2.2	4.3	7.5	0.0	0.0	0.0	0.0	0.0	2.4	15.8	19.3
Reduced Physical Function	0.0	0.0	0.0	0.1	0.2	1.6	5.5	5.7	0.0	0.0	0.0	0.3	0.6	7.7	20.9	23.6

Note: Based on RUG-III assignments that reflect 75 percent or more of a stay. This restriction was used to allow outlier results associated with a RUG-III group to reflect stays that are generally characterized by that RUG-III assignment. Results for eight RUG-III groups are based on extremely small (n<30) sample sizes: IB2, IA2, BB1, PE2, PD2, PC2, PB2, PA2.

Appendix Exhibit VIII.8: Outlier Stay and Payment Percentages under Eight NTA Outlier Payment Policies, by RUG-III Group

RUG-III Group	Outlier Stay Percents								Outlier Payment Percents							
	1%	2%	3%	4%	5%	10%	15%	16%	1%	2%	3%	4%	5%	10%	15%	16%
RUC	0.2	0.5	1.1	1.3	1.8	2.9	5.1	5.3	0.1	0.3	0.7	1.1	1.5	3.0	4.3	4.5
RUB	0.0	0.2	0.3	0.4	0.6	1.7	3.3	3.6	0.0	0.1	0.3	0.5	0.7	1.8	3.0	3.2
RUA	0.0	0.0	0.0	0.2	0.3	1.4	2.4	2.7	0.0	0.0	0.0	0.1	0.2	1.3	2.8	3.1
RVC	0.1	0.4	0.9	1.3	1.3	5.0	9.2	9.7	0.0	0.4	0.9	1.5	2.1	5.0	8.8	9.4
RVB	0.1	0.3	0.4	0.7	1.0	3.1	5.6	5.9	0.2	0.6	0.9	1.3	1.7	4.0	6.5	6.9
RVA	0.1	0.3	0.5	0.6	0.9	2.8	5.6	6.0	0.1	0.7	1.3	1.9	2.4	5.9	10.0	10.6
RHC	0.1	0.3	0.6	0.9	1.3	4.2	8.2	8.9	0.2	0.6	1.1	1.7	2.3	5.9	10.2	10.9
RHB	0.0	0.1	0.3	0.5	0.9	3.4	7.4	8.1	0.2	0.4	0.8	1.2	1.7	5.3	10.3	11.2
RHA	0.1	0.2	0.3	0.6	0.9	3.6	7.7	8.5	0.8	1.4	2.1	2.9	3.8	10.1	18.5	19.9
RMC	0.2	0.6	1.0	1.5	2.1	5.6	10.1	10.8	0.9	1.7	2.5	3.5	4.4	9.7	15.3	16.1
RMB	0.2	0.4	0.8	1.1	1.7	5.2	10.1	11.1	0.4	1.2	2.0	2.9	3.9	9.7	16.5	17.7
RMA	0.1	0.3	0.6	0.9	1.4	4.4	9.7	10.4	0.6	1.4	2.5	3.6	4.8	12.6	22.2	23.8
RLB	2.0	4.8	6.8	8.2	8.8	17.7	23.1	23.1	16.2	29.8	41.1	50.1	57.5	86.7	110.8	114.1
RLA	0.0	0.0	0.3	1.0	1.3	8.1	16.8	17.7	0.0	0.0	0.2	1.7	3.5	20.8	45.5	49.5
SE3	0.5	1.3	2.1	2.9	4.0	9.3	15.0	15.9	2.3	4.7	6.9	9.1	11.3	22.3	33.1	34.7
SE2	0.3	0.9	1.3	1.8	2.5	6.2	11.3	12.2	1.9	4.0	6.0	7.9	9.8	19.4	29.5	31.1
SE1	0.2	0.2	0.5	0.5	0.5	1.1	3.6	4.5	2.5	4.0	5.1	6.2	6.9	9.3	13.8	14.8
SSC	0.4	1.0	1.4	2.0	2.4	4.9	8.5	8.7	1.7	3.0	4.2	5.3	6.3	10.5	14.9	15.6
SSB	0.2	0.7	0.8	1.3	1.9	5.9	9.4	10.1	0.6	2.0	3.0	3.7	4.6	9.3	14.5	15.3
SSA	0.5	1.8	2.9	4.1	5.3	12.6	21.4	22.5	6.2	11.7	18.0	23.9	29.8	58.3	85.8	90.0
CC2	0.0	0.0	0.0	0.0	0.0	1.1	6.4	7.4	0.0	0.0	0.0	0.0	0.0	0.5	3.3	3.9
CC1	0.0	0.2	0.7	0.7	1.2	3.1	4.8	5.8	0.0	0.6	1.3	2.0	2.7	6.5	10.4	11.1
CB2	0.0	0.3	1.0	1.7	2.1	5.5	11.0	11.6	0.0	0.2	1.2	2.6	4.1	10.8	18.7	20.0
CB1	0.4	0.5	0.6	1.0	1.6	3.9	8.8	9.3	0.9	2.9	4.2	5.3	6.5	13.3	20.9	22.1
CA2	0.3	0.9	2.4	3.0	3.6	7.0	14.3	15.5	3.5	6.3	10.2	14.6	18.2	32.7	47.4	49.8
CA1	0.3	0.7	1.1	1.6	2.0	5.9	11.4	12.3	2.1	5.3	8.3	11.2	13.9	28.9	44.8	47.4
IB2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
IB1	0.0	0.0	0.0	0.0	0.0	0.4	3.5	4.7	0.0	0.0	0.0	0.0	0.0	0.2	4.7	6.2
IA2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
IA1	0.0	0.0	0.0	0.0	0.0	2.7	6.4	6.8	0.0	0.0	0.0	0.0	0.0	12.4	28.4	31.9
BB2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BB1	0.0	0.0	0.0	0.0	0.0	4.2	8.3	8.3	0.0	0.0	0.0	0.0	0.0	0.3	18.6	21.9
BA2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BA1	0.0	0.0	0.0	0.0	0.0	1.4	2.9	7.2	0.0	0.0	0.0	0.0	0.0	3.4	14.4	18.0
PE2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PE1	0.0	0.0	0.0	0.0	0.0	1.2	5.8	5.8	0.0	0.0	0.0	0.0	0.0	1.0	9.2	11.0
PD2	0.0	0.0	0.0	0.0	0.0	5.9	17.6	17.6	0.0	0.0	0.0	0.0	0.0	17.0	34.3	37.4
PD1	0.0	0.0	0.0	0.2	0.2	1.7	5.5	5.7	0.0	0.0	0.0	0.6	1.2	7.8	19.0	21.2
PC2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PC1	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	5.6
PB2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PB1	0.0	0.0	0.0	0.0	0.0	0.0	4.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	5.4	8.4
PA2	0.0	0.0	0.0	0.0	0.0	0.0	50.0	50.0	0.0	0.0	0.0	0.0	0.0	0.0	51.8	85.6
PA1	0.0	0.0	0.0	0.0	0.6	2.6	6.6	7.2	0.0	0.0	0.0	0.0	0.8	19.5	49.3	54.4

Note: Based on RUG-III assignments that reflect 75 percent or more of a stay. This restriction was used to allow outlier results associated with a RUG-III group to reflect stays that are generally characterized by that RUG-III assignment. Results for eight RUG-III groups are based on extremely small (n<30) sample sizes: IB2, IA2, BB1, PE2, PD2, PC2, PB2, PA2.

Appendix Exhibit VIII.9: Outlier Stay and Payment Percentages under Eight NTA Outlier Payment Policies, by Facility Characteristics

Characteristic	Number SNFs (in sample)	Number Stays (in sample)	Outlier Stay Percents 1%	2%	3%	4%	5%	10%	15%	16%	Outlier Payment Percents 1%	2%	3%	4%	5%	10%	15%	16%
All	*1,386*	*170,253*	*0.3*	*0.7*	*1.2*	*1.7*	*2.3*	*5.8*	*10.0*	*10.7*	*1.0*	*2.0*	*2.9*	*3.9*	*4.8*	*9.2*	*13.2*	*13.8*
Free Standing	1,203	126,948	0.2	0.5	0.9	1.3	1.7	4.0	6.7	7.1	0.7	1.4	2.1	2.7	3.4	6.2	8.9	9.3
Hospital Based	183	43,305	0.4	1.2	2.1	3.1	4.2	11.3	19.8	21.2	2.2	4.5	6.8	9.0	11.2	21.4	29.8	31.0
Rural	430	35,863	0.3	0.8	1.3	1.9	2.6	6.3	10.9	11.6	1.5	2.6	3.6	4.6	5.7	10.4	14.7	15.3
Urban	956	134,390	0.2	0.7	1.2	1.7	2.3	5.7	9.8	10.5	0.9	1.8	2.8	3.7	4.6	8.9	12.9	13.5
Non-profit	361	56,302	0.3	0.9	1.4	2.1	2.9	7.6	13.5	14.5	1.2	2.6	3.9	5.2	6.5	12.8	18.5	19.4
Government	80	11,033	0.3	0.9	1.8	2.5	3.6	9.7	17.2	18.2	1.6	3.1	4.8	6.5	8.1	16.3	23.4	24.4
For-profit	945	102,918	0.2	0.6	1.0	1.4	1.9	4.4	7.4	7.8	0.9	1.6	2.4	3.1	3.8	7.0	9.9	10.3
Chain	786	85,074	0.2	0.6	1.1	1.5	1.9	4.6	7.6	8.1	0.9	1.7	2.5	3.2	3.9	7.2	10.2	10.6
Non-chain	159	17,844	0.2	0.5	0.9	1.2	1.6	3.8	6.3	6.9	0.6	1.3	1.9	2.5	3.2	6.0	8.6	9.0
0-49 Medicare beds	595	73,947	0.3	0.9	1.5	2.2	3.0	7.8	13.7	14.6	1.3	2.7	4.0	5.4	6.7	13.1	19.0	19.8
50-99 Medicare beds	436	42,872	0.3	0.7	1.1	1.6	2.0	4.8	8.0	8.6	1.1	1.9	2.8	3.6	4.3	7.8	11.1	11.5
100+ Medicare beds	355	53,434	0.2	0.5	0.9	1.2	1.6	3.9	6.6	7.0	0.6	1.3	1.9	2.5	3.1	5.9	8.5	8.9
0-9% Medicare days	704	45,657	0.2	0.6	1.0	1.3	1.7	3.9	6.8	7.2	1.2	2.0	2.7	3.4	4.1	7.1	9.9	10.4
10%-15% Medicare days	372	46,284	0.2	0.6	0.9	1.3	1.7	4.0	6.9	7.3	0.5	1.2	1.8	2.4	3.0	5.8	8.5	8.9
16%-25% Medicare days	142	28,020	0.2	0.6	1.0	1.4	1.7	4.0	6.8	7.2	0.9	1.6	2.2	2.9	3.5	6.3	8.8	9.2
25%+ Medicare days	168	50,292	0.3	1.1	1.8	2.7	3.8	10.1	17.7	18.9	1.6	3.5	5.4	7.3	9.2	18.2	25.8	26.9
New England	112	12,382	0.1	0.3	0.6	0.9	1.3	3.4	6.5	7.0	0.3	0.7	1.2	1.7	2.2	4.7	7.5	7.9
Mid-Atlantic	171	29,087	0.2	0.6	1.0	1.4	1.8	4.6	8.4	9.0	0.7	1.4	2.1	2.8	3.4	6.7	9.9	10.4
South Atlantic	216	31,533	0.2	0.7	1.3	1.8	2.4	5.8	9.8	10.4	0.6	1.5	2.4	3.4	4.3	8.7	12.7	13.3
East North Central	273	34,341	0.3	0.6	1.1	1.6	2.1	5.1	8.9	9.5	1.2	2.2	3.0	3.9	4.8	8.8	12.5	13.1
East South Central	93	12,768	0.4	0.9	1.4	1.7	2.1	5.4	9.7	10.5	1.8	3.0	4.1	5.1	6.0	10.1	14.0	14.5
West North Central	155	12,486	0.2	0.7	1.2	1.8	2.5	6.3	11.3	12.1	1.6	2.7	3.8	4.8	5.9	11.3	16.2	16.9
West South Central	143	13,237	0.5	1.4	2.4	3.5	4.8	11.8	19.2	20.4	2.2	4.2	6.2	8.1	10.0	18.5	25.4	26.4
Pacific	146	16,135	0.2	0.8	1.2	1.7	2.3	5.9	9.9	10.5	0.9	2.0	3.1	4.0	5.0	9.7	14.2	14.8
Mountain	75	7,995	0.2	0.5	1.1	1.8	2.3	6.0	10.2	10.8	0.5	1.3	2.2	3.2	4.2	8.9	13.3	13.9

Appendix Exhibit VIII.10

Outlier Stay and Payment Percentages under Eight NTA Outlier Payment Policies, by Facility Characteristics (Using Alternative Loss-Share Methodology)

	Outlier Stay Percents								Outlier Payment Percents							
	1%	**2%**	**3%**	**4%**	**5%**	**10%**	**15%**	**16%**	**1%**	**2%**	**3%**	**4%**	**5%**	**10%**	**15%**	**16%**
All	*0.3*	*0.7*	*1.2*	*1.8*	*2.4*	*5.9*	*10.1*	*11.0*	*1.0*	*2.0*	*2.9*	*3.9*	*4.8*	*9.1*	*13.0*	*13.8*
Free-Standing	0.2	0.5	0.9	1.3	1.7	4.0	6.7	7.3	0.8	1.5	2.2	2.8	3.4	6.4	9.0	9.6
Hospital Based	0.4	1.2	2.1	3.1	4.3	11.4	19.9	21.8	2.1	4.3	6.5	8.7	10.8	20.6	28.7	30.2
Rural	0.3	0.8	1.3	1.9	2.6	6.4	10.9	12.0	1.5	2.6	3.6	4.6	5.6	10.3	14.4	15.3
Urban	0.2	0.7	1.2	1.7	2.3	5.8	9.9	10.7	0.9	1.8	2.8	3.7	4.6	8.8	12.7	13.5
Non-profit	0.3	0.9	1.5	2.2	3.0	7.7	13.6	14.8	1.2	2.5	3.8	5.0	6.3	12.5	18.0	19.1
Government	0.3	0.9	1.8	2.6	3.6	9.8	17.3	18.7	1.6	3.0	4.6	6.2	7.8	15.7	22.5	23.8
For-profit	0.2	0.6	1.1	1.4	1.9	4.5	7.4	8.1	0.9	1.7	2.4	3.2	3.9	7.1	10.0	10.6
Chain	0.2	0.6	1.1	1.5	1.9	4.6	7.6	8.3	0.9	1.7	2.5	3.3	4.0	7.3	10.2	10.8
Non-chain	0.2	0.5	0.9	1.2	1.6	3.9	6.4	7.0	0.6	1.3	2.0	2.6	3.2	6.1	8.7	9.2
0 - 49 Medicare beds	0.3	0.9	1.5	2.2	3.0	7.9	13.7	15.0	1.3	2.6	4.0	5.3	6.6	12.9	18.4	19.5
50 - 99 Medicare beds	0.3	0.7	1.1	1.6	2.1	4.9	8.1	8.8	1.1	2.0	2.8	3.6	4.4	7.9	11.1	11.7
100+ Medicare beds	0.2	0.5	0.9	1.3	1.6	3.9	6.6	7.2	0.6	1.3	1.9	2.6	3.2	6.0	8.5	9.1
0% - 9% Medicare days	0.3	0.6	1.0	1.3	1.7	4.0	6.8	7.5	1.2	2.0	2.7	3.4	4.1	7.2	10.0	10.6
10%-15% Medicare days	0.2	0.6	0.9	1.3	1.7	4.1	6.9	7.5	0.5	1.2	1.8	2.5	3.1	5.9	8.6	9.1
16%- 25% Medicare days	0.2	0.6	1.0	1.4	1.8	4.1	6.9	7.4	0.9	1.6	2.3	2.9	3.6	6.4	8.9	9.5
25%+ Medicare days	0.4	1.1	1.9	2.8	3.9	10.3	17.8	19.4	1.6	3.4	5.3	7.1	8.9	17.6	24.9	26.3
New England	0.1	0.3	0.6	0.9	1.3	3.4	6.6	7.3	0.3	0.7	1.2	1.7	2.2	4.8	7.5	8.1
Mid-Atlantic	0.2	0.6	1.0	1.4	1.8	4.7	8.4	9.2	0.7	1.4	2.1	2.8	3.5	6.8	9.9	10.6
South Atlantic	0.2	0.7	1.3	1.8	2.5	5.9	9.8	10.6	0.6	1.5	2.5	3.4	4.4	8.7	12.6	13.4
East North Central	0.3	0.6	1.1	1.6	2.1	5.2	8.9	9.7	1.2	2.2	3.1	4.0	4.8	8.8	12.4	13.2
East South Central	0.4	0.9	1.4	1.7	2.1	5.4	9.8	10.8	1.8	3.0	4.1	5.0	5.9	9.9	13.6	14.4
West North Central	0.3	0.7	1.2	1.8	2.5	6.5	11.3	12.4	1.6	2.7	3.7	4.8	5.9	11.1	15.8	16.7
West South Central	0.5	1.4	2.4	3.5	4.9	12.0	19.2	20.9	2.1	4.1	6.0	7.9	9.8	18.1	24.7	25.9
Pacific	0.2	0.8	1.2	1.7	2.3	6.0	10.2	11.2	0.9	2.0	3.1	4.0	5.0	9.6	13.9	14.7
Mountain	0.2	0.5	1.2	1.8	2.3	6.1	10.0	10.8	0.5	1.3	2.2	3.2	4.2	8.9	13.2	14.0

Appendix Exhibit VIII.11: Average Outlier Payments Across all Stays under Eight NTA Outlier Payment Policies, by Facility Characteristics

	Number SNFs (in sample)	Number Stays (in sample)	Average Outlier Payments per Stay							
			1%	2%	3%	4%	5%	10%	15%	16%
All	*1,386*	*170,253*	*11*	*22*	*32*	*42*	*53*	*101*	*146*	*153*
Free-standing	1,203	126,948	9	18	26	33	41	74	104	109
Hospital Based	183	43,305	16	34	51	69	87	180	268	281
Rural	430	35,863	15	26	36	47	57	105	148	155
Urban	956	134,390	10	21	31	41	51	100	145	152
Non-profit	361	56,302	11	24	37	49	62	125	185	194
Government	80	11,033	14	26	40	55	69	145	218	229
For-profit	945	102,918	10	20	29	38	46	83	117	121
Chain	786	85,074	11	21	30	39	48	86	121	126
Non-chain	159	17,844	7	15	22	30	37	69	97	101
0-49 Medicare beds	595	73,947	12	25	37	50	63	127	187	196
50 - 99 Medicare beds	436	42,872	13	23	33	42	51	92	128	133
100+ Medicare beds	355	53,434	8	16	24	32	39	73	103	108
0% - 9% Medicare days	704	45,657	13	22	31	38	46	79	109	113
10%-15% Medicare days	372	46,284	6	15	22	30	37	71	101	106
16%- 25% Medicare days	142	28,020	12	21	29	38	46	81	113	117
25%+ Medicare days	168	50,292	13	28	44	60	77	160	239	251
New England	112	12,382	4	9	14	20	26	56	87	91
Mid-Atlantic	171	29,087	8	17	26	35	43	82	120	126
South Atlantic	216	31,533	6	16	26	37	47	95	137	144
East North Central	273	34,341	14	24	33	43	52	96	135	141
East South Central	93	12,768	19	31	42	52	61	103	142	148
West North Central	155	12,486	15	24	34	44	54	104	150	157
West South Central	143	13,237	20	40	60	79	99	193	276	288
Pacific	146	16,135	10	24	37	48	59	116	170	178
Mountain	75	7,995	5	13	22	32	42	91	136	143

Appendix Exhibit VIII.12: Average Outlier Payments Across all Stays under Eight NTA Outlier Payment Policies, by Facility Characteristics (Using Alternative Loss-Share Methodology)

	Average Outlier Payments per Stay							
	1%	**2%**	**3%**	**4%**	**5%**	**10%**	**15%**	**16%**
All	*11*	*22*	*32*	*42*	*53*	*100*	*144*	*153*
Free-Standing	9	18	26	34	42	76	106	112
Hospital Based	16	32	49	67	84	172	255	271
Rural	15	26	36	46	56	104	146	154
Urban	10	21	31	41	51	100	143	152
Non-profit	11	24	36	48	60	122	179	191
Government	13	25	39	53	66	139	207	221
For-profit	11	20	30	38	47	85	118	124
Chain	11	21	31	40	49	88	122	128
Non-chain	8	16	23	31	29	71	99	105
0-49 Medicare beds	12	24	37	49	62	124	181	193
50 - 99 Medicare beds	13	23	33	43	52	93	128	136
100+ Medicare beds	8	17	25	33	40	74	104	110
0% - 9% Medicare days	13	23	31	39	47	81	110	116
10%-15% Medicare days	6	15	23	31	38	72	103	109
16%- 25% Medicare days	12	21	30	39	47	83	114	120
25%+ Medicare days	12	27	43	58	74	154	228	244
New England	4	9	15	21	27	57	87	94
Mid-Atlantic	9	18	27	35	43	83	120	128
South Atlantic	6	16	27	37	47	95	136	145
East North Central	14	24	34	43	53	96	134	142
East South Central	19	31	42	52	60	101	139	146
West North Central	14	24	34	43	53	101	145	155
West South Central	20	39	58	77	96	187	266	282
Pacific	10	24	36	48	59	115	167	177
Mountain	5	13	22	32	43	91	135	144

Appendix Exhibit VIII.13: Ratios of Total Payments to Costs under Eight NTA Outlier Payment Policies, by Facility Characteristics

	Number SNFs (in sample)	Number Stays (in sample)	Ratio of RUG-III Payments to Costs per Stay	Ratio of Total NTA Payments to Costs per Stay							
			No Outlier Policy	1%	2%	3%	4%	5%	10%	15%	16%
All	*1,386*	*170,253*	*0.95*	*0.95*	*0.95*	*0.95*	*0.95*	*0.95*	*0.95*	*0.95*	*0.95*
Free-standing	1,203	126,948	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08
Hospital Based	183	43,305	0.59	0.59	0.60	0.60	0.60	0.60	0.61	0.62	0.62
Rural	430	35,863	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
Urban	956	134,390	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96
Non-profit	361	56,302	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.79	0.79
Government	80	11,033	0.68	0.68	0.68	0.68	0.68	0.68	0.69	0.69	0.70
For-profit	945	102,918	1.09	1.09	1.09	1.09	1.08	1.08	1.08	1.08	1.08
Chain	786	85,074	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.08	1.08
Non-chain	159	17,844	1.07	1.07	1.06	1.06	1.06	1.06	1.06	1.06	1.06
0-49 Medicare beds	595	73,947	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.81	0.81
50 - 99 Medicare beds	436	42,872	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05
100+ Medicare beds	355	53,434	1.09	1.09	1.09	1.09	1.09	1.09	1.08	1.08	1.08
0% - 9% Medicare days	704	45,657	1.06	1.06	1.06	1.06	1.06	1.06	1.05	1.05	1.05
10%-15% Medicare days	372	46,284	1.09	1.09	1.09	1.09	1.09	1.09	1.08	1.08	1.08
16%- 25% Medicare days	142	28,020	1.12	1.12	1.12	1.12	1.12	1.12	1.11	1.11	1.11
25%+ Medicare days	168	50,292	0.66	0.66	0.66	0.66	0.66	0.66	0.67	0.68	0.68
New England	112	12,382	1.02	1.02	1.02	1.02	1.02	1.02	1.01	1.01	1.01
Mid-Atlantic	171	29,087	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
South Atlantic	216	31,533	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
East North Central	273	34,341	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
East South Central	93	12,768	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
West North Central	155	12,486	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86
West South Central	143	13,237	0.80	0.80	0.81	0.81	0.81	0.81	0.82	0.82	0.83
Pacific	146	16,135	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
Mountain	75	7,995	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97