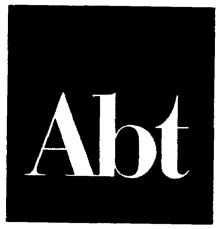


Abt Associates Inc.

55 Wheeler Street
Cambridge, Massachusetts
02138-1168

617 492-7100 *telephone*
617 492-5219 *facsimile*

Hampden Square, Suite 600
4800 Montgomery Lane
Bethesda, Maryland
20814-5341

301 913-0500 *telephone*
301 652-3618 *facsimile*

640 North LaSalle Street
Suite 400
Chicago, Illinois
60610-3781

312 867-4000 *telephone*
312 867-4200 *facsimile*

Variation in Prescribed Medication Costs, Method of Collection and Impact on Skilled Nursing Facility Casemix

Technical Expert Panel Briefing Materials

Preliminary Findings

February 28, 2000

Not for quotation, reproduction or distribution

Sponsored by:
Department of Health and Human Services
Health Care Financing Administration

Prepared by:
Abt Associates Inc.
55 Wheeler Street
Cambridge, MA 02138

Principal Authors:
Brant E. Fries
Kate Lapane
Terry Moore
Vincent Mor
Steve Pizer
Alan White

Table of Contents

Executive Summary 1

1.0 Overview 6

2.0 Data Sources, Creation of Cost and Analytic Measures 9
2.1 Data Sources 9
2.2 Creation of Cost Measures 10

3.0 Descriptive Analyses 15

4.0 Methodology 24
4.1 Overview 24
4.2 Identification of Variables and Exploration of Interactions 24
4.3 Index Model Construction 25
4.4 Test and Validation Samples 26
4.5 Evaluating Potential Refinements 26

5.0 Development of Potential Refinements 31
5.1 Relationship Between RUG-III and Costs 31
5.2 Ability of RUG-III to Predict Ancillary Charges and Total Costs 34
5.3 Costs for Residents Who Qualify for Both Extensive Services and Rehabilitation 35
5.4 MDS Items Associated With Differences in Ancillary Charges 37
5.5 Clinical Review and Modification 40
5.6 Description of Potential RUG-III Refinements 45

References 52

Appendix A The RUG-III Classification System
Appendix B Minimum Data Set (MDS) Version 2.0
Appendix C Analysis of MDS Items Associated With Differences in Prescription Drug, Respiratory Therapy, or Other Ancillary Charges

List of Tables

Table 3.1 Sociodemographic Characteristics of Residents of SNF Stays by RUG-III Group 18
Table 3.2 Active Clinical Diagnoses of Residents of SNF Stays by RUG-III Group ... 19
Table 3.3 Active Clinical Diagnoses of Residents of SNF Stays by RUG-III Group ... 20
Table 5.1 Distribution of Ancillary Charges by RUG-III Category..... 33
Table 5.2 Predictive Power of RUG-III Classification System 35
Table 5.3 Comparison of Total Costs for Those in Rehabilitation Categories Based on Whether Resident Also Qualifies for Extensive Services..... 37
Table 5.4 MDS Items Associated With Higher Total Ancillary Charges 39
Table 5.5 MDS Items Associated With Differences in Ancillary Charges—Refined Variable List Following Clinical Input 44
Table 5.6 Statistical Performance of Potential RUG-III Refinements—Model Description..... 49

List of Figures

Figure 3.1 23

Executive Summary

The Health Care Financing Administration (HCFA) implemented, under Congressional mandate, a prospective payment system (PPS) for Medicare skilled nursing facility (SNF) care. This PPS categorizes SNF residents by care need, using the Resource Utilization Groups, Version 3 (RUG-III), and became effective in July 1998.

Despite general agreement that casemix-adjusted prospective payment systems furnish desirable incentives for efficiency and reduce administrative burden for both providers of services and for payers, concern has been expressed about the ability of this particular payment system - the Medicare SNF PPS - to capture adequately variance in certain types of costs. RUG-III more accurately captures variance in the *staff and therapy* resources used to care for SNF residents than any other patient classification system developed to date. However, the classification system by design did not take into consideration other types of resources that contribute to care of the SNF resident, such as prescription medications, oxygen and other non-therapy ancillary supplies and services. To allay concerns that access to quality SNF care may become restricted for Medicare beneficiaries with high non-therapy ancillary costs, and in recognition that such charges now comprise about one quarter of the daily costs of care of the Medicare SNF resident, HCFA awarded this contract to Abt Associates and our partners, Brown University Center for Gerontology and Health Care Research and the University of Michigan Institute of Gerontology to evaluate potential improvements to RUG-III to be used in the PPS for Medicare Skilled Nursing facility care.

The goal of this study is to review the RUG-III classification system with particular emphasis on the care needs of medically complex Medicare beneficiaries and the variation in non-therapy ancillary services within RUG-III categories.

A key part of this research was the exploration of potential refinements to the Extensive Services category. Previous research showed that the Extensive category is associated with the highest per diem non-therapy ancillary costs of any of the RUG-III categories. The research also indicated that, while the Extensive Services category did capture a disproportionate share of high cost beneficiaries, there was considerable within-group variance in costs. In this project, additional studies were conducted to extend the analysis of non-therapy ancillary costs and within-group variance to other RUG-III categories.

The research focused on the following analyses to identify options, and the results were used to develop the proposed RUG-III refinements discussed in this report:

1. Evaluate the ability of the current RUG-III system to predict variance in drug, respiratory or other non-therapy ancillary costs.

2. Evaluate the ability of specific MDS items to predict variance in non-therapy ancillary costs, and identify the MDS items most closely associated with differences in non-therapy ancillary costs.

3. Design/test potential refinements to the RUG-III methodology.

Using a research database comprised of Nursing Home Minimum Data Sets (MDS) matched to Medicare SNF claims, we performed a series of analytic steps to develop and evaluate potential refinements to RUG-III. These began with a search for variables that are associated with large differences in costs for residents, either overall or within RUG-III categories. Through an iterative process, we identified a group of MDS items that are associated with differences in prescription drug, respiratory therapy, or other non-therapy ancillary charges. In this search, we did not limit ourselves only to those variables already utilized in the RUG-III classification system. We also looked at RUG-III modifications, for example addressing potential interactions between existing RUG-III categories (such the Rehabilitation and Extensive Services categories).

A variety of options were designed that redefined existing RUG groups, added new terminal end-splits, and created new groups that could be added to the current 44-group system. Each refinement was evaluated using statistical, clinical, incentive, and administrative considerations. A Clinical Workgroup was convened in November 1999 to assist the project team in selecting those MDS variables considered appropriate for inclusion in any modified casemix system.

In our development of a casemix system to predict non-therapy ancillary costs, we considered a large variety of approaches. During the analysis, however, several guiding principles emerged that helped focus our work:

- *Maintain the integrity of the RUG-III System.* We wanted to augment, but not replace the RUG-III system. The system has been proven in a large number of studies, both in the US and abroad, to be effective in explaining nursing and therapy costs, and has been adopted by Medicare for reimbursement of SNF services. We consider two changes: 1) form an extra hierarchy category for those residents who could qualify for both the Rehabilitation and Extensive Services groups; and 2) extend the system with additional, "leafy end" splits that increase the explanation of ancillary costs.

- *Modification only to the "Upper" RUG-III Hierarchy.* We found that there was little to be gained by modification of the RUG-III Impaired, Behavior, or Physical Functions categories. This was due to a combination of low utilization of drugs and other non-therapy ancillary services, and low predictive capability of MDS variables to explain these costs. With only very modest gains in explanatory power achievable, we chose not to focus on modifying or altering this part of the RUG-III system.

- *"Leafy End" splits.* We sought to maintain not only the basic RUG-III structure, but also its basis in a tree-based classification. Thus, the modifications we sought were primarily designed as splits of the RUG-III groups.

- *Index-based Models.* Analytic results pointed to the use of regression-derived index models of multiple MDS variables rather than AID-derived interactive tree structures based on "indicator variables." The use of multiple variables has a disadvantage of adding complexity to the system, but the advantage that increased payment is likely linked to more than one single resident characteristic or service.

- *Focus on total ancillary charges.* Our experience with casemix analysis of composite costs has been that understanding the predictors of individual components of this composite has significant analytic advantage. For example, in the derivation of an earlier RUG-T18 system for Medicare (Fries et al., 1989), confusing results for total cost were understood when we analyzed separately its two components: nursing and therapy costs (which were negatively correlated, causing the complexity). Here we found both some overlap and lack of overlap among the predictors of individual ancillary costs, but little interaction among and relatively low predictability of several of these components. Thus, after identifying predictors of each component, we used these predictors in unified regressions of total cost.

- *New categories for residents who qualify for Extensive Services and a RUG-III rehabilitation category.* All of the index model-based refinements include the addition of five new categories: one for each of the five Rehabilitation categories (Ultra-high Rehabilitation, Very-high Rehabilitation, High Rehabilitation, Medium Rehabilitation, and Low Rehabilitation) who also qualify for the Extensive Services category. These categories would go at the "top" of the casemix system (i.e., residents would be considered first for inclusion in these categories); the Extensive Services and the existing Rehabilitation categories which would now include only one or the other. The ADL splits for these five new categories would be the same as for the corresponding old Rehabilitation category.

RUG-III refinements which we recommend for consideration by HCFA and the Technical Expert Panel include:

- *Model RUG-III+*: This is the RUG-III model with new categories for residents who qualify for Extensive Services and one of the RUG-III Rehabilitation categories, as described above. The resulting casemix system would have 14 additional Extensive Services and Rehabilitation Groups, which would use the same Rehabilitation categories and ADL splits as the current Rehabilitation categories.[1]

- *Model WIM 1:* Applies a weighted index model to Extensive Services residents (including residents in the new Extensive Services and Rehabilitation categories). A disproportionate share of high cost residents qualify for Extensive Services, so it may make sense to apply the index model only to those residents. Using this refinement, the casemix system would have up to 143 groups if the index model were incorporated within RUG-III as new terminal splits. Alternatively, the system could be implemented as a six-group ancillary add-on system for the Extensive Services residents.[2]

- *Model WIM 2:* Applies a weighted index model to Extensive Services residents (including residents in the Extensive Services and Rehabilitation category), and to Rehabilitation, Special

1 It should be noted that RUG-III+ could also be implemented as a new terminal split within the existing Rehabilitation categories based on whether the resident also qualified for Extensive Services. Statistically, the two systems are identical.

2 The preliminary weighted index model relies on a somewhat subjective determination of how to discriminate among the different index levels: i.e., estimates of the relative differences between ancillary cost levels associated with the various MDS items.

Care, and Clinically Complex residents. In this model, there would be up to 258 groups if it were implemented as new terminal splits rather than as a 6-group ancillary add-on system for these categories.

- *Model UWIM:* Applies an unweighted index model to Extensive Services residents (including residents in the Extensive Services and Rehabilitation category), and to Rehabilitation, Special Care, and Clinically Complex residents. In this model, there would be up to 178 groups if it were implemented as new terminal splits rather than as a 4-group ancillary add-on system for those categories.

A graphic representation of each of these models may be found in the appendix.

Performance of the proposed models was evaluated through a variety of measures, including statistical performance (R-squared, sensitivity and specificity), clinical coherence, and administrative complexity (e.g., number of groups). For our purposes, R-squared is a statistic that measures how much of the variance in costs observed in the data can be explained or predicted by the alternate RUG-III model. A brief summary of the performance of each of the potential refinements follows.

RUG-III+: Adding the new Extensive Services and Rehabilitation categories resulted in small improvements in statistical performance. The validation sample R-squared increased to 7 percent for ancillary charges, an increase of about 4 percent relative to RUG-III, and to 12.3 percent for total costs. Predicted costs for the 58 groups in the RUG-III+ model ranged from $104 to $384 (for one of the new Extensive Services and Ultra-high Rehabilitation groups).

WIM 1: Application of WIM 1 resulted in some improvement relative to RUG-III+. For the validation sample, the model accounted for 9 percent of the variance in ancillary charges and 16 percent of the variance in total costs. Nearly 29 percent of residents in the top 10 percent of ancillary charges were also in the top 10 percent in terms of predicted costs, compared to 22 percent for RUG-III. Under WIM1, Extensive Services residents (including those in the new Extensive Services and Rehabilitation categories) would receive an ancillary "add-on" based on the index model variables applicable to the resident. A six group ancillary index was used. There would be no additional ancillary "add-on" for residents whose predicted costs are below the 50th percentile in terms of predicted ancillary charges, $17 for those between the 51st and 74th percentile, $34 for those in the 75th - 89th percentile, $56 for those in the 90th - 94 th percentile, $106 for those in the 95th - 98th percentile, and $225 for those in the top one percent.

WIM 2: Model WIM2 was the most statistically powerful refinement that we examined. The WIM2 model accounted for 20 percent of the variance in total costs and 14 percent of the variance in ancillary charges. The range of payments was similar to that of WIM1. Using WIM2, 32 percent of residents in the top 10 percent in terms of actual ancillary charges were also in the top 10 percent in terms of predicted ancillary charges. The model applies a six-group ancillary index to 40 RUG-III+ groups (14 Rehabilitation/Extensive Services groups, 3 Extensive Services groups, 14 Rehabilitation groups, 3 Special Care and 6 Clinically Complex groups), and results in a large number of groups if it is implemented as part of an integrated classification system. Alternatively, as with the other index model-based refinements, WIM2 could be thought of as a

six group ancillary add-on which works alongside RUG-III to determine total payment. Statistically, the two systems are identical.

UWIM: This model is the unweighted counterpart to WIM2. While this model performed better than the RUG-III and RUG-III+ models, it did not perform as well as WIM2. UWIM accounted for 11 percent of the variance in ancillary charges and 18 percent of the variance in total costs. The range of payments for UWIM was quite similar to that of the weighted index models. The sensitivity and specificity of the model were slightly less than for WIM2. Using UWIM, residents are split into four groups based on the *number* of index model variables present. The splits used were 0 (45 percent of test sample observations), 1-2 (45 percent), 3-5 (9 percent) and 6 or more (0.4 percent). Residents with no index model items present would receive no ancillary payment, while those with 1-2 items present would receive $19 (based on predicted ancillary charges), those with 3-5 items would receive $68 and those with six or more would receive $209.

1.0 Overview

Background

Among the payment reforms mandated by the Balanced Budget Act (BBA) of 1997 (H.R. 2015) is the requirement that the Department of Health and Human Services implement a prospective payment system (PPS) for Medicare skilled nursing facility (SNF) care. Implementation of this system began on July 1, 1998. The PPS is based on an all-inclusive, prospectively set *per diem* payment rate that covers all Medicare-covered services (routine, ancillary, and capital-related) provided to residents while in a Part A SNF stay, including services such as nursing care, rehabilitation therapy, pharmaceuticals, and laboratory services.

The prospective payment is casemix adjusted based on the Resource Utilization Group, Version 3 (RUG-III) resident classification system that is based on data from the Minimum Data Set (MDS) (see Appendices A and B). RUG-III, developed as part of the multi-state Nursing Home Casemix and Quality demonstration, is a 44 group casemix classification system designed to capture the resource use (staff time) of nursing home residents (Fries et al., 1994). This classification system measures the intensity of care and services required for different types of SNF residents and translates this into a payment rate.

Despite agreement that PPS furnishes desirable incentives for efficiency, there is concern that payment rates may not be suitably adjusted to the care needs of residents. RUG-III more accurately captures variance in the staff time resources used to care for SNF residents than any other classification system developed to date, but RUG-III was developed based on nursing and therapy time and may not reflect differences in medical conditions or other resident characteristics associated with higher ancillary charges (e.g., prescription drugs, medical equipment and supplies, IV therapy). In the current study, we found that the system accounted for 4 percent of the variance in *per diem* ancillary charges and 10 percent of the variance in total costs (including a simulated staff time cost measure). This finding was consistent with an earlier study which found that RUG-III accounted for only 9 percent of the variance in total costs and 7 percent of the variance in ancillary charges (Abt Associates, 1999, unpublished).

Although the casemix system cannot directly affect prescribing patterns, a system that is sensitive to the wide variations in cost associated with drug treatment may be a necessary condition for optimal drug therapy treatment to occur. If the casemix system does not offer an adequate payment rate for some types of residents, for example those who require more medically complex care or expensive prescription medications, then the PPS system may have implications for access to or quality of SNF care. Nursing facilities have some discretion over which residents to admit, and residents for whom the expected costs of care are greater than the prospective payment rate may have difficulties obtaining access to SNF services or may not receive all of the services that they need as facilities attempt to provide care within the payment amount.

Given the potential for inappropriate drug prescribing patterns and other potentially adverse outcomes to occur under national PPS, further research is needed to ensure that the payment system adequately reimburses SNFs for costs incurred for resident care.

Purpose and Goals of Study

The purpose of this study to review the RUG-III classification system with particular emphasis on the care needs of medically complex Medicare beneficiaries and the variation in non-therapy ancillary services within RUG-III categories.

The task of designing potential refinements involved several analyses:

- We examined potential refinements that were based on internal changes to RUG-III (e.g., interactions between existing categories). The most promising of these potential modifications was to create new categories for residents who qualified for both Extensive Services and a Rehabilitation category.

- We also examined potential refinements using other MDS items that were associated with higher costs. The process of identifying clinically appropriate items associated with cost differences that could be used in potential refinements involved several steps:

  – We identified the subset of MDS items that are associated with differences in prescription drug, respiratory therapy, or other non-therapy ancillary charges. These items were selected by testing a large number of variables to identify the subset with a significant relationship to costs.

  – Using MDS items associated with significant differences in either prescription drug, respiratory therapy, and other non-therapy ancillary charges, we identified items that were associated with differences in total ancillary charges.

  – Some items, despite their ability to identify high cost residents, were rejected outright due to potential negative incentive effects. Others were found acceptable, with modification, and the remainder were recommended as is for inclusion in a potential model. Clinical input for the study has come both from Abt nursing staff and from the Clinical Work Group that was assembled to review the variables underlying the index models.

Overview of recommended refinements

A number of potential types of refinements were considered, the most promising of which fell into two general categories:

- **Changes to the casemix system for residents who qualify for both Extensive Services and a rehabilitation category.** Ancillary charges for residents in the Extensive Services categories were much higher than for other residents, including those in the RUG-III rehabilitation categories. Costs were much higher for residents who qualified for both Extensive Services and Rehabilitation than for those who qualified for Rehabilitation only. These high costs suggest that, at a minimum, the payment rate for Extensive Services should be increased. Increasing the payment rate without further adjustments, however, could reduce provider incentives to provide therapy to Extensive Services residents. A new category for

residents who qualify for Extensive Services and Rehabilitation would alleviate these concerns.

- **Refinements to the casemix system based on index models.** Specific options for incorporating the MDS items that survived clinical review were developed. Analytic results pointed to the use of regression-derived index models of multiple MDS variables rather than AID-derived interactive tree structures based on "indicator variables." The use of multiple variables has a disadvantage of complexity, but the advantage that increased payment is likely linked to more than a single characteristic or service. Refinements based on index models achieve potentially important improvements in statistical performance and allow for much higher payment rates for residents with characteristics associated with high ancillary charges, including most Extensive Services residents.

2.0 Data Sources, Creation of Cost and Analytic Measures

2.1 Data Sources

Researchers at Brown University, the University of Michigan Assessment Archive Project (UMAAP), and Abt Associates assembled a large cross-linked data system including nursing home resident assessments collected using the federally-mandated Minimum Data Set (MDS), drug information, HCFA claims data, and organizational data on nursing home providers. Each of the datasets used to derive the analytic files is briefly described below.

Minimum Data Set (MDS) including drug data

The MDS care assessment component of the Resident Assessment Instrument (RAI) has over 300 data elements. Trained clinical nursing and social work staff responsible for the resident completed the MDS. Topics covered in the MDS include cognitive function, communication/hearing problems, physical functioning, continence, psychosocial well-being, mood state, activity and recreation, disease diagnoses, health conditions, nutritional status, oral/dental status, skin condition, special treatments, and medication use. In one study evaluating the entire MDS, researchers reported that 89 percent of the MDS items achieved an intra-class correlation of 0.4 or higher with 63 percent achieving an estimate of at least 0.6. For the current project, we have included data from Kansas, Maine, Mississippi, New York, Ohio, South Dakota, and Texas. Unless limited by data availability (e.g., Texas 1997 only), we included MDS data from 1995 through 1997.

As part of the MDS in each of the states included in this study, staff coded up to eighteen drugs taken within the seven days preceding the assessment. Nursing home staff coded each drug according to the National Drug Coding (NDC) system using either the 10,000 NDCs included in the MDS+ manual or the Physicians' Desk Reference Book. NDCs are unique 10-digit codes that identify drug products. We matched the NDC codes to MediSpan™ which includes over 150,000 generic drug products, products from regional manufacturers, and information on over 90,000 inactive drugs. The overall match rate between the NDC and the MediSpan™ was greater than 90 percent with only 5.4 percent of the original NDC codes recorded on the MDS Section U in the MDS+ Casemix Demonstration states being incomplete or incorrect. Gambassi et al (1997) have shown that the MDS drug data are consistent and reliable.

Health Care Financing Administration (HCFA) Claims Data

We have merged various HCFA data to the MDS files using the Health Insurance Claim number of Medicare beneficiaries. We also use gender and date of birth in the matching process. To ensure confidentiality, these identifiers have been replaced with unique identifiers using the HIC number as a seed. SNF services are a Part A (hospital insurance) benefit under Medicare and are available only to patients who require continued skilled nursing care and/or skilled rehabilitation services on a daily basis following a hospital stay of at least 3 days.

On-line Certification Automated Survey (OSCAR) Data

The OSCAR data provided nursing-home level information and was merged to the MDS file by facility code. The OSCAR data include results from survey inspections. Currently, the match rates are over 90 percent. Of particular interest from the OSCAR database is the variable documenting whether the facility is hospital-based.

2.2 Creation of Cost Measures

Because the measurement of cost is both difficult and central to the analysis of potential casemix refinements, we employed two complimentary approaches. First, we follow previous studies by building measures of non-therapy ancillary cost based on Medicare SNF claims. Second, we develop a new way to measure costs associated with prescription drug therapies by exploiting data recorded in Section U of the MDS. The difference between these approaches lies in the measurement of drug costs, since in both cases other non-therapy ancillary costs are measured by claims. Since the Section U-based measures were still being refined when the RUG-III refinement analyses reported here were conducted, the results presented in this briefing manual rely exclusively on claims. We were able to use the Section U data to describe drug utilization for the current report (see Chapter 3).

Finally, to provide an additional perspective on the performance of potential refinements, we use Staff Time Measurement Study data to impute staff time costs for each observation. It should be emphasized that, because it is imputed, staff time cost is only included in one set of analyses.

Creation of Measure of Non-therapy Ancillary Charges from SNF Claims

Non-therapy ancillary services include diagnostic services, pharmaceuticals and necessary supplies and equipment. Medicare part A SNF claims was used to measure the per-diem ancillary charges. For ancillary charges that are developed using Medicare claims data, it was not possible to identify items with a date of service that corresponds to the period covered by the MDS assessment. Per diem charges were calculated using Medicare claims with a covered date within a specified range of a date covered by MDS assessment. Operationally, per diem charge are given by the sum of the costs of the ancillary therapies divided by the number of days covered by claims. Since this procedure leads to an overestimate of the true level of the reimbursable costs, adjustments were made in the cost calculation.

We estimated the costs of non-therapy ancillaries using revenue codes as extracted from the claims data. Using revenue codes as determined by Abt Associates for another project, we identified target revenue codes and categorized costs into these conceptually meaningful categories. The categories and their related revenue codes included the following: prescription drugs/pharmacy (250-259), drugs requiring ID (630-639), IV therapy (260-269), medical and surgical supplies (270-270;620-622), respiratory services (410-419), laboratory (300-309), oxygen (600-604), dialysis (820-829, 830-839, 880-889).

It is important to note that the actual ancillary *costs* for residents in the sample are not observed. The covered charges reported in claims are routinely discounted by the intermediary responsible for processing on the basis of audited reasonable cost. Inclusion of ancillary charges without further

adjustment in our measure of *per diem* ancillary charges would overstate the true level of reimbursable costs since these charges are routinely discounted before payment under the present system.

Discount rates are computed by cost center in the process of completing the annual SNF cost report and then applied throughout the year. Unfortunately, we were not able to reproduce these calculations exactly because of missing data in the SNF Cost Report Minimum Data Set. To be as consistent as possible with this practice, we calculated one average discount factor (the ratio of total Part A allowed cost to total Part A charges) for each facility in each year. Since some facilities did not have cost reports on file for every year, we excluded from the sample facilities lacking cost reports in two or more years and calculated an average ratio across years for all remaining facilities to improve precision and fill-in any missing years. This average discount factor was applied to all the facility's non-therapy ancillary charges before analysis.

This method adjusts ancillary charges downward for most residents (i.e., those residents at facilities where the total Part A allowed cost was less than the total Part A charges), so that the contribution of non-therapy ancillaries to total costs is not overstated. We are limited in our ability to create an actual measure of resident-specific cost due to available data sources. Cost report data are facility-specific, while claims data reflect charges to the Medicare program per SNF resident. While the method that we employ for these analyses uses all of the information that is available for converting ancillary charges into a measure of ancillary costs (i.e., cost report and SNF claims data), it relies on a facility-specific adjustment factor, as again, there are no data for creating a resident-specific adjustment. We therefore use the term "ancillary charges" throughout this document to refer to the adjusted estimate of non-therapy ancillary costs.

Creation of MDS-based Drug Cost Measures (Section U)

We used the average wholesale price (AWP) as included in the MediSpan™ for medication costs for several reasons. The AWP is a national figure and not subject to regional influences resulting from purchasing contracts and other local market factors. This helps to account for the cost of dispensing. Using AWP is conservative when the price of a medication is relatively low or high AWP is not subject to institutional cost-shifting. We also evaluated differences between the pricing options for a sub-set of representative and frequently used medications. Additionally, owing to the amount of completeness with each pricing option, we used AWP because it also yielded the lowest amount of missing cost data.

While we were able to successfully map NDC codes to drug names (nested within therapeutic classes and sub-classes), to successfully match to a drug cost required more information. Specifically, to assign an AWP to a drug, both the strength of the drug administered and complete information regarding the frequency with which the medication was administered is required. Unfortunately, many of the codes included in the MDS training manual itself do not include information regarding strength. For example, we may know that a resident received aspirin, but we do not know if it was 80 mg, 325 mg, or some other strength. As a result, we have substantial missing cost data. Because of the extent of missing data, we opted to impute the drug costs as opposed to excluding cases for which

we did not have complete drug cost information. Analyses of the extent of missing data revealed that missing data did not vary by RUG group, state, year, and type of medication.

Nonetheless, by imputing missing drug costs, we have introduced random variations in the data that were not generated by the underlying process that we are attempting to model. Consequently, variables that explain variance in non-missing data will have no explanatory power for imputed data. As a result, the coefficients on these variables will be biased toward zero. This bias will be small if the proportion of total variance attributable to imputation is small. However, variables explicitly or implicitly used in the imputation process may have explanatory power with regard to the imputed values. For example, if the RUG group is implicitly used as part of the imputation process, it theoretically could as a result explain more of the variance in the dependent variable simply because RUG was used as part of the imputation algorithm. The coefficients of the variables used to impute cost data may be amplified relative to other coefficients in the explanatory models. Depending on correlation between the RUG groups and other variables, these coefficients will also be biased in unpredictable ways. This problem could be small if the between-group variance is small (overall variance can be decomposed into between-group and within-group components). Yet, if the proportion of variance attributable to imputation is small because the actual imputed values are small relative to whatever component of cost we are modeling. Given the potential for introducing bias in our models, we opted to create two imputation algorithms. The values imputed from each method were remarkably similar and did not alter any of the findings. Therefore, we discuss one of the imputation methods below and only show the analyses using this method.

Because of our concerns regarding bias, we implemented an algorithm to estimate the drug costs based on data contained in Section U of the MDS. We thought that missing data might vary systematically by state owing to differing data collection procedures (and software) used by states. Furthermore, we considered that coding of drugs might have improved through time. If both assumptions were true, the pattern of missingness would vary systematically through time and place. It follows that an imputation method based on time and place would be reasonable. If the NDC code was not listed among the 150,000 Medispan codes, but the exact name of the generic drug was listed, pricing was calculated as follows. If only one cost was associated with the drug within a given state and year, it was used. If multiple costs were associated, a cost was chosen probabilistically based on the distribution of observed costs among residents. If the exact generic name could not be matched, a match for the leading words in the generic name was made, and if a match we applied the same approach (i.e. opting to a probabilistically selected drug cost using the state and year). In cases where no reasonable match could be found, no price was assigned to the medication. As with the RUG-based imputation measure, this algorithm was iterative over the observed distribution among residents.

Exclusions and Creation of Analytic Sample

Creation of the analytic sample used to develop and test potential refinements was guided by the desire to have a large, representative sample and the need to exclude assessments likely to contain reporting errors. Our original sample included 733,300 MDS assessments from seven states, representing the years 1995-1997. We then reduced this sample through implementation of the following exclusion criteria:

1. Exclude all assessments from New York. All assessments from New York were excluded from analyses that used Medicare claims data because many facilities in the state billed SNF stays using an all-inclusive rate. Because these facilities did not use the revenue codes that we used to measure prescription drug, respiratory therapy or other non-therapy ancillary charges, measured ancillary charges for most New York residents were zero in some or all of the revenue codes analyzed for this study. The exclusion of New York results in the removal of 525,215 of the 733,300 total MDS assessments from our analytic sample.

2. Exclude all assessments for which a cost-to-charge ratio could not be calculated. Medicare cost report data were used to calculate the facility-specific ratio of Total Part A allowed cost to total Part A charges for each facility in each year. Facilities missing Medicare cost reports for at least two years between 1995 and 1997 were excluded because we were not able to calculate cost-to-charge ratios for the facility. This resulted in the exclusion of 93,314 additional assessments.

3. Exclude all facilities for which the correlation between a measure of drug costs calculated from Section U and one calculated from Medicare claims data was less than zero. We used drug charge data derived from Medicare claims in the refinement analyses, but used the Section U data to identify facilities with unreliable drug cost data. For facilities that have a negative correlation between the two drug cost measures, there is a concern about inaccurate reporting on either claims or MDS assessments at the facility level, and these facilities were excluded. This step resulted in the exclusion of 10,915 MDS assessments.

4. Exclude all residents with per diem ancillary charges greater than $1,000. Two hundred fifty-three (253) observations with *per diem* total ancillary charges greater than $1,000 were excluded from the refinement analyses. Summary measures of statistical performance such as R-squared are typically sensitive to outliers, and these extreme values were judged unlikely to be accurate. In addition, such values have disproportionate leverage in the design of potential refinements. The exclusion of extreme outliers in refinement analyses does not mean that their costs cannot be considered when determining payment rates.

The resulting analytic sample included 103,856 assessments, which were assigned randomly to either the test or validation samples. We assigned 60 percent of this sample– 61,929 assessments– to the test sample which was used to develop and test potential refinements. The remaining 41,927 assessments comprised the validation sample.

Limitations to Cost Measurement

There are several limitations to the cost measures created, including problems with the relationship between the MDS assessment observation period, the usual claim covered period, and the derivation of the cost to charge ratio. All of these factors mean that a direct assignment of Medicare costs per sampled resident is not possible. The ancillary charges measures calculated from SNF claims thus have two important potential sources of bias.

- *Claims generally cover a different, typically longer, period than the MDS assessment.* It was not possible to identify the day that the specific services included in Part A SNF claims were

received and *per diem* estimates are based on the average costs across the entire period covered by the claim.

- *The charge amount that appears on the claim did not necessarily reflect what HCFA actually paid to the SNF for the service.* These costs were routinely discounted before payment under the fee-for-service system. We used the adjustment factor described above to correct for this discrepancy. The adjustment factor ensured that, in the aggregate, non-therapy ancillary charges are an appropriate proportion of total observed costs, but there are no data which would allow us to calculate resident- or service-specific adjustment factors. As a result, there are doubtless differences between the non-therapy ancillary charge measures that we used and what HCFA actually paid for the service. It is not possible to determine how large these differences were, but they may have been quite large in some cases.

Measurement error in the calculation of non-therapy ancillary charges reduces our ability to model the sources of variance in non-therapy ancillary charges, reducing the measured accuracy of the casemix system. Potential measurement error has several implications for our analyses. The difference in the dates covered by the Medicare claim and the MDS assessment results in a less precise estimate of the relationship between MDS items and non-therapy ancillary charges. This affects measurements of the statistical performance of both RUG-III and potential refinements that are based on the MDS. The imprecision with which non-therapy ancillary charges were measured introduces a source of variation in costs that cannot be captured by either RUG-III or the potential refinements, and almost certainly results in an underestimate of the predictive power of all of the models discussed in this report, including RUG-III. It is not possible to estimate the size of this underestimate.

3.0 Descriptive Analyses

As mentioned in Chapter 2, the analyses described here were supported by a variety of data sources, including Minimum Data Set data, SNF claims data, cost report, staff time measurement and OSCAR data. Though we were unable to use MDS Section U data for our RUG refinement analyses, the data do offer insights about the study population, and the variation in drug utilization by RUG-III category. This chapter discusses descriptive findings using MDS Section U data.

Characteristics of the sample

Table 3.1 shows the sociodemographic characteristics of the sample stratified by an aggregate of the RUG categories. The majority of residents were female (65 percent), with little variation in the proportion across the RUG categories. Residents classified as having only behavior problems were less likely to be male (37 percent) and those with reduced physical functioning were the least likely to be male (30 percent). The majority of residents were white, of non-Hispanic origin (84 percent). Approximately nine percent of residents were Black and 2 percent were Hispanic. Overall, nearly one quarter of residents was severely cognitively impaired. Among residents classified in a Rehabilitation RUG category, 35 percent were moderately impaired and 14 percent were severely cognitively impaired. The distribution of cognitive impairment among those classified as Physical Function Reduced was similar to that of the Rehabilitation RUG category. Residents classified as Extensive Services or Special Care also had a similar distribution of cognitive impairment level. Approximately one third of each were moderately impaired. Thirty-nine percent of residents were classified as dependent in activities of daily living and only 7 percent with no limitations. Residents categorized as having only Behavioral Problems were most likely to have only minimal limitations in physical functioning (28 percent). Residents classified as Clinically Complex (14 percent), Cognitively Impaired (13 percent), or Reduced Physical Functioning (14 percent) were also more likely to have minimal limitations relative to the other RUG categories. Residents in the Extensive Services (58 percent) and Special Care (56 percent) categories were most likely to be classified as dependent in activities of daily living.

The active clinical diagnoses documented for residents in the sample are shown stratified by RUG group on Table 3.2. Cardiovascular diseases were common in residents. Overall 20 percent of residents had coronary artery disease. Cardiac arrhythmia was present in 14 percent of residents. Overall, nearly one quarter of residents had congestive heart failure and 9 percent had peripheral vascular diseases. On average, 43 percent of residents had documented hypertension. While the distribution of residents with coronary artery disease appeared similar across RUG groups, congestive heart failure and arrhythmia were more common in the Extensive Services, Special Care, and Clinically Complex categories. For most of the cardiovascular conditions, residents in the Impaired Cognition category were less likely to have these diseases relative to other RUG categories. A similar but attenuated pattern was noted for residents in the Behavior Only RUG group.

Neurological diseases were also common. Overall 9 percent of residents had Alzheimer's disease documented. Twenty-eight percent had other dementia documented. Nearly one quarter of residents had an active clinical diagnosis of stroke and 6 percent had Parkinson's disease. While the proportion

of residents with Parkinson's disease did not vary by RUG group, the proportion with other neurological conditions varied substantially by RUG group. Residents in the Impaired Cognition group were more likely to have Alzheimer's disease (22 percent) and other dementia (54 percent) documented and the less likely to have had a stroke (15 percent) compared to other RUG groups. Similar to the Impaired Cognition group, residents in the Behavior Problem category were more likely to have other dementia (41 percent) and less likely to have had a stroke (12 percent) compared to other RUG groups, but had a similar proportion of residents with Alzheimer's disease. The distribution of neurological conditions among residents classified as Extensive Services, Special Care, and Clinically Complex were similar to distributions of the former two. A third of residents classified as Extensive Services and Special Care had non-Alzheimer's dementia and one quarter had suffered a stroke.

Only 5 percent of residents had anxiety and 16 percent had depression documented as a diagnosis on the MDS. Across RUG groups, the proportion of residents with anxiety and depression was similar. However, the prevalence of anxiety (8 percent) and depression (22 percent) was higher in the Behavior Problem RUG category. Twelve percent of residents had cataract and 7 percent glaucoma. These conditions did not vary substantially by RUG group. Overall, septicemia was rare (1 percent), and only 8 percent of residents had pneumonia and 17 percent had urinary tract infections. Residents in the Extensive Services categories were more likely to have septicemia (2 percent), pneumonia (17 percent), and urinary tract infections (24 percent) compared to other RUG categories. Other diagnoses and conditions were common. Twenty-one percent of residents had allergies, 19 percent anemia, 22 percent had arthritis, 22 percent had diabetes, and 12 percent had cancer. Residents in the Rehabilitation, Extensive Services, Special Care, and Clinically Complex categories were more likely to have these conditions relative to the Impaired Cognition and Behavioral Problem RUG categories. The prevalence of hypothyroidism (10 percent) did not vary by RUG group.

Pooling across all states and the three years, there is little variation by RUG group in total daily drug cost as measured by Section U. Median costs within the Rehabilitation RUG groups range from ~$6.50 (Low Rehab groups) to ~$9.00 (Ultra-high Rehab groups) whereas the lowest costs of medications was experienced by the Impaired Cognition groups (~$3.00). The groups with the higher interquartile range (~$13) were the Extensive Services categories and some of the rehabilitation RUG groups (e.g., RVC ~$12). The Impaired Cognition groups also demonstrated the least variation in costs of medications with an interquartile range of ~$5.

To better understand which classes of drugs may be driving costs, we classified the drugs according to fourteen major therapeutic classes. The most expensive therapeutic drug classes are anti-infective agents (Median: $6.53) and biologics (Median: $9.73). The least expensive therapeutic drug classes are analgesics (Median: $0.10) and nutritional products (Median: $0.18). The proportion of residents within each of the major RUG categories are shown in Table 3.3. Variations in medication use across RUG groups were apparent for many medication classes and corresponded to observed variations in the active clinical diagnoses shown by RUG group in Table 3.2. Residents were least likely to be on biologics (1 percent) and anti-neoplastics (2 percent), regardless of RUG class. The majority of residents were on at least one cardiovascular medication, with substantial variation across RUG groups. Residents in the Rehabilitation RUG groups (67 percent) and in the Clinically Complex group (64 percent) were the most likely to be receiving at least one cardiovascular medication. Residents in

the Impaired Cognition (47 percent) and Behavioral Problem group (53 percent) were the least likely to be receiving cardiovascular medications.

Similar trends were observed across RUG groups for both gastrointestinal agents and endocrine/metabolic agents. More than half of residents had taken at least one gastrointestinal agent with residents in the Rehabilitation categories (67 percent) the most likely to use gastrointestinal products and residents in the Impaired Cognition or Behavioral Problem RUG groups the least likely to receive these drugs (~50 percent). With endocrine and metabolic agents, over one third of residents of Rehabilitation, Extensive Services, Special Care, and Clinically Complex RUG groups received these drugs, relative to ~25 percent of other RUG groups. Residents in the Rehabilitation, Extensive Services, Special Services, and Clinically Complex RUG groups were most likely to be on anti-infective agents, with over 25 percent of residents in each on these medications. Amongst these RUG groups, residents of Table 3.1

Table 3.1

Sociodemographic Characteristics of Residents of SNF Stays by RUG-III Group.

	All	Rehabilitation	Extensive Services	Special Care	Clinically Complex	Impaired Cognition	Behaviors only	Physical Function Reduced
Male	35	37	36	34	36	35	37	30
Race/Ethnicity								
White	84	90	83	83	82	80	84	83
Hispanic	2	1	2	2	2	3	3	2
Black	9	6	9	9	9	11	8	9
Asian/Pacific Islander	0.5	0.2	0.7	0.5	0.6	0.7	0.7	0.6
American Indian	1	0.7	2	2	2	1	1	1
Missing†	3	.9	3	4	4	3	3	3
Cognitive Impairment:✣								
Mild (CPS: 0-1)	41	51	33	35	47	0	50	53
Moderate (CPS: 2-4)	35	35	31	34	35	67	50	32
Severe (CPS: 5-6)	23	14	34	31	17	33	0	14
Physical Functioning:								
Minimal limitations	7	6	0	3	14	13	28	14
Moderate limitations	44	53	37	36	51	58	49	47
Dependent	39	18	58	56	31	20	7	26
Missing†	9	23	6	4	4	9	16	12

✣ CPS = Cognitive Performance Scale

†Missing data percentages shown when greater than 3% missing data occurred.

Totals may not equal 100% due to rounding.

Table 3.2

Active Clinical Diagnoses of Residents of SNF Stays by RUG-III Group

	All	Rehabilitation	Extensive Services	Special Care	Clinically Complex	Impaired Cognition	Behaviors only	Physical Function Reduced
Heart/Circulation								
Coronary artery disease	20	14	22	22	22	21	19	21
Cardiac arrhythmia	14	15	16	15	14	11	8	12
Congestive heart failure	24	22	27	25	27	16	20	21
Hypertension	43	44	42	42	44	37	40	42
Peripheral vascular diseases	9	8	10	12	9	6	7	7
Other cardiovascular diseases	20	20	21	21	21	16	16	17
Neurological								
Alzheimer's disease	9	5	9	9	8	22	11	8
Other dementia	28	18	30	30	27	54	41	28
Cerebrovascular disease	23	26	24	25	25	15	12	16
Parkinson's disease	6	5	6	6	5	6	5	6
Psychiatric								
Anxiety	5	6	5	5	6	5	8	5
Depression	16	17	15	17	18	15	22	15
Sensory								
Cataract	12	6	14	14	14	14	13	13
Glaucoma	7	5	7	7	7	6	8	7
Infections								
Septicemia	1	1	2	1	1	0	0	0
Pneumonia	8	8	17	8	10	0	0	0
Urinary tract infection	17	16	24	19	13	10	9	12
Other								
Allergies	21	23	22	22	21	14	19	17
Anemia	19	16	23	22	19	15	14	17
Arthritis	22	22	23	22	21	17	19	24
Cancer	12	11	14	13	13	7	8	9
Emphysema/COPD	15	14	17	15	19	10	14	10
Diabetes mellitus	22	22	22	23	24	15	19	18
Hypothyroidism	10	10	10	10	10	9	9	9
Osteoporosis	8	9	8	8	8	6	6	9

Table 3.3

Active Clinical Diagnoses of Residents of SNF Stays by RUG-III Group

	All	Rehabilitation	Extensive Services	Special Care	Clinically Complex	Impaired Cognition	Behaviors only	Physical Function Reduced
Anti-infectives	26	29	39	28	23	12	12	16
Biologics	1	0.3	1	2	1	1	1	1
Anti-neoplastics	2	2	2	2	3	1	2	1
Endocrine	31	36	30	30	33	22	26	26
Cardiovascular	61	67	59	59	64	51	55	58
Respiratory	19	23	21	18	23	9	17	13
Gastrointestinal	61	67	60	62	62	47	53	58
Genitourinary	5	6	5	5	5	4	3	5
CNS	36	43	32	33	38	46	55	34
Analgesics	47	60	43	45	44	32	39	44
Neuromuscular	13	13	13	13	12	14	18	12
Hematological	30	35	30	31	29	20	19	26
Topical	30	26	34	37	28	20	20	23

the Extensive services categories were the most likely to be taking anti-infective agents (39 percent). Less than 15 percent of residents in other RUG groups received these drugs.

Overall, 47 percent received at least one analgesic. Impaired Cognition Residents (32 percent) and Behavior Problem Residents (39 percent) were less likely to receive analgesics than those in the Rehabilitation group (60 percent). Similar trends were apparent with hematological agents (~20 percent Impaired Cognition vs. ~35 percent in the Rehabilitation groups), and topical agents (~20 percent vs. ~37 percent in the Special Care groups). Conversely, residents in the Impaired Cognition (~46 percent) and Behavioral Problem (over 50 percent) RUG categories were more likely to receive CNS drugs relative to the other RUG groups (~33 percent).

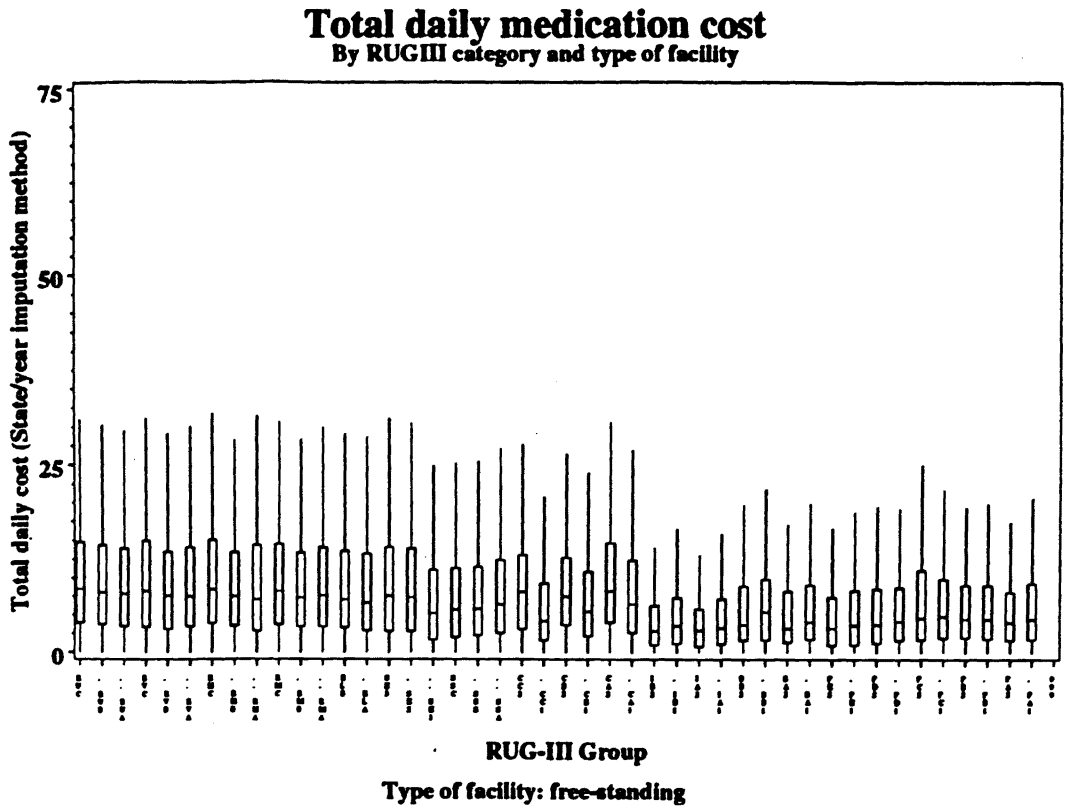
The highest proportion of total costs due to anti-infective use is found in the Extensive Services and Clinically Complex RUG groups with ~ 50 percent of drug costs attributable to the anti-infective agents. Use of biologics was relatively infrequent (~1.2 percent) and the proportion of drug costs due to these agents was highly variable amongst the users, regardless of RUG group. Among people receiving anti-neoplastic medications(~2.2 percent of residents), these agents accounted for one quarter of their total daily drug cost (Median: 27 percent; 25th percentile: 13 percent; 75th percentile: 49 percent). Regardless of RUG group, this measure is highly variable. While nearly one third of all residents received an endocrine medication, these agents only accounted for 8 percent of the total daily drug costs amongst users. Cardiovascular medications accounted for 18 percent of the total daily drug cost, which varies slightly across RUG group (+/- ~4 percent). There appears to be slightly less variation in this measure among the Extensive Services, Special Care, and Clinically Complex groups as compared to other RUG categories. Among the 19 percent of residents using respiratory medications, 12 percent of their drug costs were due to these agents. Higher median proportions and greater variability occurred at the end splits within the aggregate RUG categories. A similar pattern is observed among users of gastrointestinal agents. These medications accounted for only 13 percent (Median) of the total daily costs. This measure is highly variable, regardless of RUG group. Only 5 percent of residents had used a genitourinary medication, accounting for only 13 percent of total drug costs (Median value). This measure varied slightly across RUG group.

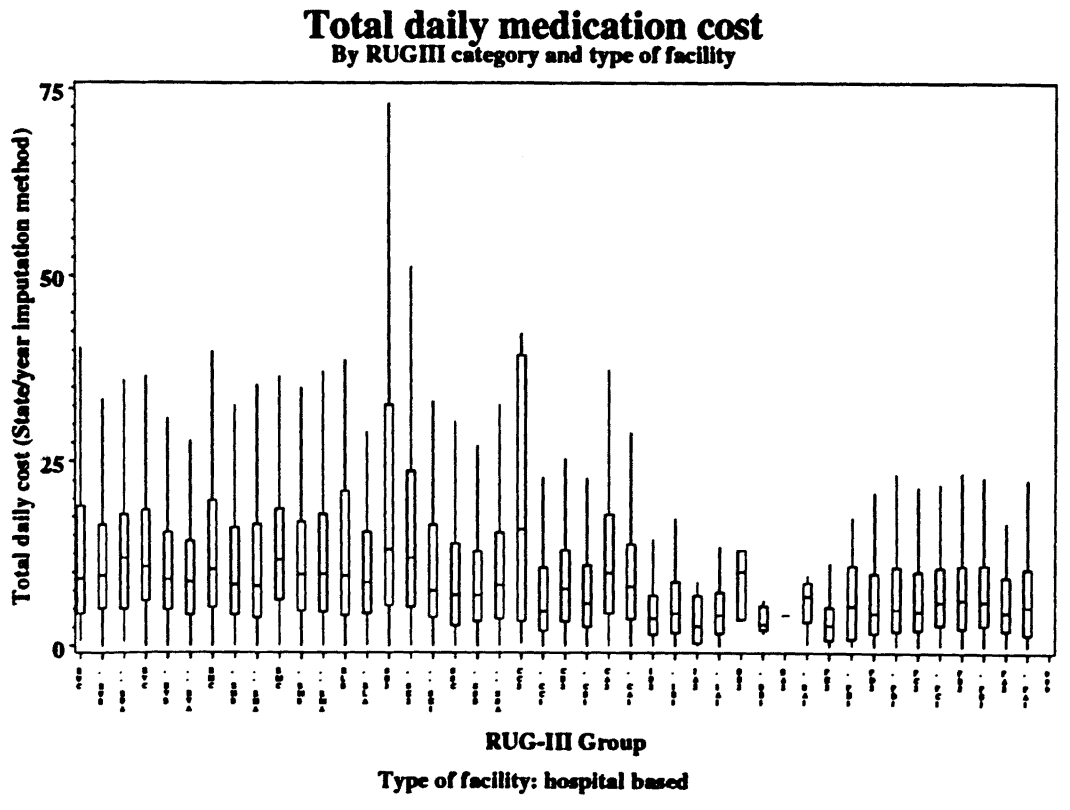
Contrasting Hospital Based and Free-Standing SNF Drug Costs

Historically, SNF care in hospital based nursing facilities has been reimbursed at a higher rate than care provided in free-standing skilled nursing facilities. Part of this cost difference has been attributed to the mix of patients both in terms of their nursing and therapy needs. Presumably, risk adjustment using RUGs grouping should account for this difference. Residual differences that persist are attributed to the fact that hospital based SNF patients present with more complex medical care needs. One component of that is the array of drugs and their associated costs that patients in hospital based facilities versus freestanding facilities are taking. Since PPS reimbursement does not differentiate between patients as a function of where they are served, it is pertinent to examine the extent to which there are differences in the estimated drug costs of patients in these two kinds of facilities. Since all "costs" have been standardized to the Average Wholesale Price, the existence of systematic differences in pricing between the two types of facilities have been removed and the only differences should be related to the relative costliness of the drugs similar patients in the two types of facilities receive.

Figure 3.1 (Panel A and B) presents the total daily drug costs by RUG class for patients in hospital based and free-standing facilities. Looking at the free-standing based panel it is evident that there is little variation in total daily drug cost by RUG group. The Behavior RUGs have the lowest drug costs (median between $8 and $12) whereas the Extensive and the Clinically Complex groups (CA1 and CA2) have costs that are about twice as high. Looking at the hospital based panel, there are several points that are immediately obvious. First, almost all RUG classes have higher degrees of variability than is the case for the free-standing panel. Second, Extensive, Special Care and Clinically Complex classes all have higher costs than do their counterparts in the free-standing panel. Third, a not insignificant number of Extensive and Clinically Complex patients have daily average wholesale drug costs well in excess of $25 and many have costs in excess of $50. Since these differences are not attributable to different pricing patterns across the two types of facilities, it is clear that patients in hospital based facilities are being treated with a more expensive profile of medications than are their free-standing counter-parts.

Figure 3.1


Total daily medication cost
By RUGIII category and type of facility
Total daily cost (State/year imputation method)
75
50
25
0
RUG-III Group
Type of facility: free-standing


Total daily medication cost
By RUGIII category and type of facility
Total daily cost (State/year imputation method)
75
50
25
0
RUG-III Group
Type of facility: hospital based

4.0 Methodology

4.1 Overview

The goal of this analysis is to improve RUG-III's predictive power while incorporating clinical and other factors. Regardless of the type of potential refinement being considered, the process of identifying possible refinements to RUG-III involved searching for variables that are associated with large differences in costs for residents (either overall or within RUG-III categories). The process was an iterative one that involved testing a large number of variables to identify the subset with a significant relationship with costs. For purposes of developing casemix refinements, the ideal type of variable is one that meets all of the following criteria: (1) has a significant impact on costs, (2) makes clinical and administrative sense, and (3) is observed in a large number of residents.

Once such variables were identified, a variety of refinements were designed that applied these variables to redefine existing RUG groups, add new terminal end-splits, and create new "add-on" groups that could be added to the current 44-group system. The project team developed several alternative refinements, all based on the same MDS variables, then evaluated each refinement with reference to statistical, clinical, incentive, and administrative considerations.

As stated earlier, we used claim-based cost measures, although they did not correspond to the MDS period (claims typically cover a one month period). This measure of drug costs was created by converting the charges that appear on the claims to a best guess of actual costs using an adjustment that is based on the total charges and total costs reported on the facility's cost report.

4.2 Identification of Variables and Exploration of Interactions

The current RUG-III system was developed based only on staff time costs and accounts for a relatively small proportion of the variance in non-therapy ancillary charges. Therefore, to supplement RUG-III, we examined potential refinements based on MDS items that are associated with higher non-therapy ancillary charges. These MDS items were identified by a broad and systematic search of all items in the MDS. The identified items included a variety of measures of resident acuity and treatments received, including items related to functional status (Sections G and H of the MDS), disease diagnoses (Section I), health conditions (Section J), nutritional status (Section K), skin conditions (Section M), and special treatments and procedures (Section P of the MDS).

Having assembled the effective variables, we then explored the utility in ADI-based branching models (using PC-Group), much as the RUG-III system was developed. Although the details are not reported here, models were considered both for the full population and for subsets including those classified into the Extensive category and in combined Rehabilitation and Extensive categories – these latter because of earlier results that these categories represented many of the high-cost residents.

Tree-based splitting such as this has a tactical advantage when there are strong statistical interactions present: when one characteristic is important in explaining the dependent variable of interest only in

the presence of a second (or more) characteristic(s). A hypothetical example would be that diabetes was only predictive of ancillary costs when there were pressure ulcers present; perhaps for those without ulcers, a measure of physical functioning would be predictive. Following this example, if we found that diabetes was a useful splitting variable in both branches — i.e., regardless whether the resident had pressure ulcers — then these two variables are not interactive.

In our analysis of ancillary costs, the results did not indicate strong interaction effects. There were two implications of this finding. First, the variables' effects were principally additive and models which develop indexes are indicated. Second, the appropriate approach was to use regression analysis to form indices, rather than PC-Group to identify tree models. (It should be noted that PC-Group still has some unique capabilities, employed later, to help identify optimal thresholds for an index.)

One way an index model could be used is in an "add-on" system for predicting non-therapy ancillary charges. RUG-III could be used for predicting staff time costs and a non-therapy ancillary index would be "added-on" to determine the total payment rate for residents with given characteristics. The motivation for this approach is that RUG-III has been well tested and validated for predicting staff time costs, but does not perform as well in capturing variance in non-therapy ancillary charges. Although such a system can be described as consisting of two components, it could easily be implemented as an integrated system, as though the non-therapy ancillary component defined a new set of end-splits to RUG-III.

4.3 Index Model Construction

The index model approach allowed for a large number of items to be considered simultaneously in determining payment rates, including additional measures of severity that are not reflected in RUG-III. We designed both weighted and unweighted versions of a non-therapy ancillary index and showed that both versions resulted in large improvements in the proportion of the variance predicted by the casemix system and some improvement in the system's ability to identify high-cost residents. The weighted version allowed items that predict much higher costs (such as pneumonia) to have more impact on predicted costs than less-influential items (e.g., pressure ulcer). Consequently, the weighted index model exhibited enhanced explanatory power, but at the cost of additional complexity.

The steps used to identify MDS items and create the index models are as follows:

- Examining each MDS variable independently, we identified all MDS items that have a significant positive relationship (at the 5 percent level) with *per diem* non-therapy ancillary charges, using t-tests for binary variables and univariate regression analyses for continuous measures.

- For variables that were found to be significantly related to *per diem* costs in the first step, we estimated a backward stepwise regression to identify the subset of items that in a multivariate context were still related to costs at the 5 percent level.

- We reviewed the surviving variables to evaluate their clinical validity and potential incentive effects if included in the payment rate. For example, indwelling catheters and other MDS

items that may be quality-of-care indicators were removed from consideration as casemix adjustors due to the potential incentive factors introduced. Establishing a higher payment rate for residents with these types of treatments or conditions might result in a casemix system that induces an increase in the proportion of residents with these conditions.

- Once variables were identified, a weighted non-therapy ancillary charge index score was calculated for each resident. The index score was based on how many of the selected variables apply to the resident, weighted by the importance of the variable in predicting ancillary charges. These weights were obtained as coefficients estimated from an ordinary least squares (OLS) regression of non-therapy ancillary charges on the list of selected variables, conducted on the test sample. For the unweighted model, each variable was used with a constant weight of 1.

- Finally, residents were grouped according to their weighted or unweighted index score. For both versions, splits were defined based on analysis of the test sample and applied to the validation sample for evaluation.

4.4 Test and Validation Samples

The recursive strategies employed by stepwise regression, AID, and other fitting techniques may produce over-optimistic measures of variance explanation. For that reason, assessment of the explanatory power of alternative models required use of data that were not used in forming the models themselves. We selected at random three-fifths of the sample for use as a test sample and the remaining two-fifths for use as a validation sample. Refinements to RUG-III were developed based solely on analysis of the test sample and evaluated solely on their performance with the validation sample. Since aberrations in the test sample that may have influenced the design of refinements were absent in the validation sample, any unsupported features of the proposed models should be exposed by this approach.

4.5 Evaluating Potential Refinements

The primary purpose of the resident classification system was to predict costs accurately, while providing incentives to furnish appropriate care and to classify residents into groups that made clinical sense. Evaluation of potential refinements to RUG-III is a complicated process that often involves tradeoffs between the statistical, clinical, incentive-related and administrative factors that must be considered in assessing alternative casemix systems. For example, statistical performance (in terms of the percentage variance explanation) is often maximized by the use of measures based on the *provision* rather than the *need* for services. Such measures, however, are often subject to gaming or upcoding, and may give providers the incentive to alter their practice patterns. The tradeoff between statistical performance and the avoidance of unwanted incentives is an inescapable outcome of the limited ability of any assessment measures yet developed to predict residents' needs based solely on health conditions.

The criteria used to evaluate potential refinements to RUG-III are described below. The potential refinements discussed in this report increased the statistical performance of the casemix system and met with the approval of clinical consultants and project team members. However, this increased statistical performance in general was achieved at the cost of greater complexity to the system.

Statistical Factors

If one were to graph each resident's costs on a graph with total cost on one axis and casemix categories on the other, a perfect classification system would look like a straight line with a positive slope, reflecting greater resource use for residents in higher categories. In practice, such a relationship will never be observed due to both the intrinsic variability in resource use even by residents with the same observable characteristics, and the complex relationships among staffing, payment, resident needs, and the provision of services. In addition, measurement error and unobservable resident characteristics reduce the ability of any classification system to produce such a relationship.

Despite these limitations, a casemix system that accounts for a substantial proportion of the underlying variance in expected costs reduces the financial risk to providers and also reduces the scope for skimming of financially attractive residents (i.e., those for whom the prospective payment is much greater than the expected costs of providing care). The fairness of the casemix system (to providers, beneficiaries, and the government) is enhanced by maximizing the variation in expected costs captured by the system.

R-squared is a statistic that measures how close a particular classification system comes to the ideal. This statistic is estimated routinely and reported by most statistical software as part of ordinary least squares (OLS) regression output. In the context of our models, R-squared is a measure of how much of the variance in resource use observed in the data can be explained or predicted by the model. It gives the percentage of the variation of the dependent variable (cost) explained linearly by variation in independent variables (casemix groups). Formally, this equals the sum of squared deviations of the predicted values of the dependent variables about their mean (i.e., the explained variation from the OLS regression) divided by the total variation of the dependent variable about its mean (the total sum of squares).

Since the classification system could have a relatively high variance explanation, but still fail to account for some high cost residents, potential refinements were also evaluated based on reductions in the proportions of residents whose costs of care were much higher than the payment rate. These outlier analyses were intended to measure the extent to which the overall R-squared of the model was disproportionately affected by residents with outlier costs, and to allow us to measure the proportion of residents who may experience difficulties under PPS in obtaining access to SNF services or in receiving all needed services.

Using a constructed measure of total costs (which includes *per diem* imputed staff time costs and charges for non-therapy ancillaries), we evaluated each potential refinement to RUG-III using several measures of statistical performance:

- **R-squared on the test and validation samples**
  The R-squared on the validation sample is a realistic measure of how well the model would perform in the real world. It is expected that the R-squared on the validation sample will be somewhat lower than the R-squared on the test sample. A significant drop from the test sample to the validation sample indicates one of several possibilities: the validation sample, even though chosen by chance, may contain more outliers or more extreme outliers than the test sample; or the coefficients generated using the test sample may be unreliable. This could be caused by the observations being divided into too many categories, by inadequate sample size or because coefficients are being estimated based on outlier values.

- **Maximum/Minimum group costs**: This measure provided the highest and lowest mean ancillary cost across all the payment groups. A system with a greater range is more likely to be acceptable to the industry as it will provide situations where high levels of payment will be authorized.

- **Specificity and Sensitivity in identifying high-cost residents**
  An alternative way to evaluate how a casemix system accounts for the needs of heavy care residents is to think of the system like a diagnostic test. With what probability will the casemix system predict that a resident has costs above the 90th percentile when their actual costs are above the 90th percentile? This probability is known as "sensitivity." A very sensitive system, however, may be likely to classify too many residents as high cost. Hence we also ask with what probability will the system predict that a resident has costs below the 90th percentile when their actual costs are below the 90th percentile. This probability is known as "specificity." A good system will exhibit both high sensitivity and high specificity.

Clinical Factors

In addition to adjusting for differences in costs associated with the expected resource requirements imposed by residents with different needs, the classification system needs to make sense clinically. The clinical relevance of the system is enhanced if the definitions used to classify groups within the casemix system include residents who are similar not only in terms of costs, but also in terms of medical conditions and physical and functional status.

The categories in the RUG-III system were defined to ensure that residents within each category had clinical affinity (see Fries et al., 1994, Schneider et al., 1991, Cornelius et al.,1994), based on input from a broad-based clinical panel.

Because the potential casemix refinements evaluated for this study took as given the general structure of the RUG-III system, the clinical input that went into the design of RUG-III was reflected in the potential refinements described in this report. In addition, input on the clinical appropriateness of the MDS items that qualified for the index models was also considered. Items that had a relationship with costs that could not be reconciled with clinical meaningfulness were excluded from the model.

The clinical and statistical factors used to evaluate possible casemix refinements occasionally conflict. For example, the statistical performance of measures based on the presence of specific types of disease

diagnoses used to define casemix categories or end splits has been disappointing, as most of the diagnoses have little relation to observed costs. Diagnosis-based measures may, however, improve the clinical meaningfulness of the casemix system by increasing the clinical affinity between patients in given categories. A strong case can be made for incorporating diagnosis-based measures regardless of the effects of their inclusion on the statistical performance of the system based on such clinical criteria.

In the design of possible refinements to RUG-III, we have attempted to avoid the use of measures that do not make sense from a clinical viewpoint. Clinical input for the study has come both from Abt nursing staff and from the Clinical Work Group that was assembled to review the MDS variables underlying the index models.

Incentive-related Factors

The casemix system may create incentives for providers to alter their practice patterns, or their assessment of resident needs, in a manner that will produce increased payment. The criteria used to define the categories in the casemix system will affect provider incentives, and these incentive-factors must be considered in evaluating possible refinements to RUG-III.

The classification system should not give facilities the opportunity to "game" the system by developing (or upcoding) a characteristic at little cost to the facility that results in increased payment. To minimize upcoding, categories should be as broad and inclusive as possible, subject to the statistical and clinical criteria described above. An exception to this general rule can be made for variables that, while subject to gaming, provide appropriate incentives, such as the provision of rehabilitation therapies.

Subtleties in the elements that make up the classification system can send strong incentives to providers. Making the casemix payment contingent on the types of services received can lead to much higher utilization of those services. As a result, wherever possible, measures used in the classification system should be based on measures of the *need* for the service rather than the *provision* of the service itself. Casemix systems that are based on the *receipt* of services tend to be more accurate than those relying on measures of the *need* for services, so there is often a tradeoff between improving the statistical performance of the classification system and the potentially adverse incentive effects introduced by the use of service-based measures.

There are elements both in RUG-III, and in some of the potential refinements to RUG-III, that are based on the actual utilization of services. The RUG-III rehabilitation categories are defined based on the amount and types of therapy received by residents. Several of the variables that were most strongly related to *per diem* ancillary charges and included in the index models are based on special treatments and procedures received by the resident (from Section P of the MDS 2.0).

Following Schneider et al. (1991), service-based variables should be avoided unless they meet at least some of the following criteria:

- The cost of providing a procedure of set of services offsets a significant portion of the increased payment.
- The service requires medical authorization and there is the potential of negative consequences to the resident if it is inappropriately provided.
- There are no measures available in the MDS which predict the need for the service.

In order to mitigate any inappropriate incentives created by the inclusion of service-based variables, we linked several of these items to specific diagnoses or conditions. The requirement that a service variable be linked with a clinical variable should help to limit the application of these services to the subset of residents for whom such services could be clinically appropriate.

5.0 Development of Potential Refinements

The primary goal of our analyses was to develop potential refinements that improve the ability of the casemix system to account for variation in ancillary charges, while incorporating clinical criteria and concerns. Development of potential refinements began with a set of analyses that evaluated the relationship between ancillary charges and the current RUG-III system. These analyses indicated that the predictive power of the casemix system could be improved if it included new categories for residents who qualified for both Extensive Services and a Rehabilitation category.

We also examined potential refinements using other MDS items that were associated with higher costs, using the methods described in Chapter 4. These analyses involved identifying the subset of clinically appropriate MDS items associated with differences in ancillary charges, reviewing the resulting list of variables for clinical appropriateness, and considering alternative ways of incorporating these items into a refined casemix system.

Note that prescription drug cost measures for these refinement analyses are based on Medicare claims data, rather than MDS Section U. This is because the Section U data had not yet been finalized at the time that the refinement analyses were performed. Our preliminary analyses of the MDS items associated with the measure of total ancillary charges were similar regardless of whether Medicare claims or Section U data were used to measure drug costs, as were the statistical performance of RUG-III and potential refinements.

5.1 Relationship Between RUG-III and Costs

Ancillary charges were much higher for Extensive Services residents than for residents in other RUG-III categories. Across the other categories, ancillary charges were higher for the Rehabilitation, Special Care, and Clinically Complex categories than for the Impaired Cognition, Behavior Problems, and Reduced Physical Functioning categories.

- Mean prescription drug charges were nearly $24 per day. They were more that twice as high for Extensive Services residents ($46) than for any other category except Ultra-high Rehabilitation. There were relatively small differences in mean drug charges across the other categories, which ranged from $14 for Impaired Cognition to $23 for Ultra-high Rehabilitation (Table 5.1). We also examined the distribution of residents with high drug charges by RUG-III category. Among those in the top one percent in terms of total drug charges, 36 percent were in the Extensive Services category.

- Respiratory therapy charges were highest for Extensive Services residents, although the difference was not as large as for prescription drugs. Mean respiratory therapy charges were $25 for Extensive Services residents, $9 for Special Care, and $14 for those in the Clinically Complex category. Among Rehabilitation residents, there was a strong relationship between respiratory therapy charges and the Rehabilitation category for which residents qualified. The distribution of respiratory therapy charges was highly skewed.

Less than 13 percent of the sample had any respiratory therapy charges. Among those with non-zero charges, the mean of costs was $108 and the standard deviation was $103.

- Charges for other non-therapy ancillary charges (which includes medical and surgical supplies, IV therapy, laboratory, blood, and other miscellaneous ancillary charges) for Extensive Services residents were $24, nearly three times higher than for any other category other than Ultra-high Rehabilitation. Furthermore, a disproportionate share of residents with high other ancillary charges were in the Extensive Services category. Among those with other ancillary charges of $100 or higher (the top 1.3 percent of the sample), 37 percent were in Extensive Services.

While the Extensive Services category contained a disproportionate share of residents with high ancillary charges, there was considerable within-category variation in ancillary charges. The standard deviation of ancillary charges for Extensive Services residents was $152, more than 1.5 times higher than the mean. One goal in developing potential refinements is to reduce the within-category variance of these ancillary charges.

Table 5.1

Distribution of Ancillary Charges by RUG-III Category

RUG-III category	N	Mean total ancillary charges (std. dev.)	Mean drug charges (std. dev)	Mean respiratory therapy charges (std. dev.)	Mean other ancillary charges (std. dev.)
All	61,929	45.84 (89.82)	14.27 (52.42)	23.82 (50.06)	7.75 (36.30)
Ultra-high rehabilitation	5,321	56.42 (103.92)	24.58 (66.46)	23.22 (51.91)	8.62 (36.68)
Very-high rehabilitation	5,121	49.29 (93.75)	19.14 (60.23)	22.72 (47.56)	7.43 (37.73)
High rehabilitation	4,548	40.56 (84.26)	13.32 (52.18)	21.87 (47.39)	5.38 (26.29)
Medium rehabilitation	13,523	41.36 (80.56)	13.20 (50.02)	21.86 (41.41)	6.29 (29.73)
Low rehabilitation	1,112	29.73 (60.50)	6.62 (39.30)	19.08 (33.81)	4.02 (14.17)
Extensive Services	5,525	95.49 (152.07)	25.19 (73.50)	45.91 (90.93)	24.40 (79.67)
Special Care	13,508	38.95 (71.25)	9.35 (39.03)	22.47 (45.83)	7.12 (26.98)
Clinically complex	8,086	38.65 (71.80)	13.39 (51.42)	20.95 (37.08)	4.31 (19.46)
Impaired Cognition	1,016	22.14 (44.91)	5.54 (31.38)	14.82 (26.33)	1.78 (9.15)
Behavior problems	126	27.86 (60.17)	10.68 (48.58)	15.65 (18.00)	1.53 (5.43)
Reduced physical functioning	3,986	28.11 (57.93)	6.79 (34.92)	17.94 (33.43)	3.38 (24.05)

Notes: N=61,929 (Based on test sample only)

Data Source: Medicare MDS and SNF Claims Data 1995-1997

5.2 Ability of RUG-III to Predict Ancillary Charges and Total Costs

To test the ability of RUG-III to predict ancillary charges and a measure of total costs, we estimated a regression of costs on a set of binary indicator variables for each RUG-III group (with a single group omitted to serve as the reference category). Two types of models were tested – one using all 44 RUG-III groups and the second using the "first" 26 groups (i.e., from Rehabilitation through Clinically Complex), which include most Medicare-covered SNF residents.

R-squared is a statistic that measures how close a particular classification system comes to the ideal. In the context of our models, R-squared is a measure of how much of the variance in costs observed in the data can be explained or predicted by the model. RUG-III accounted for only 10 percent of the variance in (simulated) total costs and 4 percent of the variance in ancillary charges (Table 5.2). The statistical performance was slightly lower in the model that included only residents through Clinically Complex. Examining ancillary costs separately by type, RUG-III accounted for 2 percent of the variance in prescription drug charges, about 1 percent of the variance in respiratory therapy charges, and less than one percent of the variance in other non-therapy ancillary charges.

Sensitivity of Statistical Performance to Outliers Depending on how costs are distributed, it is possible that small numbers of outliers can cause a large decrease in the R-squared of a regression of costs on casemix groups. Extreme outliers (defined as those with total ancillary charges of $1,000 or higher) were excluded from our analyses. To test how model results were affected by outliers, we estimated the regression models described above excluding all residents with ancillary charges of $100 or more (roughly the top 10 percent). With this group excluded, RUG-III predicted 18 percent of the variance in total costs (including simulated staff time costs). The R-squared of ancillary charges fell from 4 to 1 percent. The decrease in statistical performance for ancillary charges was due to the disproportionate share of high cost residents in the Extensive Services category.

Even considering the limitations in our creation of a measure of ancillary charges (see Chapter 2), these findings raise concerns about the adequacy of RUG-III in reflecting variance in non-therapy ancillary charges.

Table 5.2

Predictive Power of RUG-III Classification System

	R-squared (r^2)		
	All RUG-III Categories	**All RUG-III Categories through Clinically Complex**	**Exclude high cost residents (> $100)**
Total cost (includes simulated staff time costs)	10.0%	7.8%	18.3%
Total ancillary charges	4.1%	3.8%	1.4%
Drug charges	2.2%	2.3%	0.007%
Respiratory therapy charges	1.3%	1.2%	0.004%
Other non-therapy ancillary charges	0.007%	0.007%	0.01%

Sample size: 61,788 for "All RUG-III Categories" ; 56,672 for "All RUG-III Categories through Clinically Complex"; 54,578 for model that excludes high cost residents.

Sources: Medicare claims and MDS assessment data 1995-1997.

5.3 Costs for Residents Who Qualify for Both Extensive Services and Rehabilitation

Under the current PPS system, the payment rate (under an index maximization approach) is the same for residents who qualify for both Extensive Services and one of the top three Rehabilitation categories (Ultra High, Very High and High Rehabilitation) as for residents who qualify only for one of the top three Rehabilitation categories.

Ancillary charges were much higher for residents who qualified for both Extensive Services and a Rehabilitation category than for those who qualified only for a rehabilitation category. Across all Rehabilitation categories, mean ancillary charges were $119 for those who also qualified for the Extensive Services category and $37 for those who qualified only for Rehabilitation (Table 5.3). Large differences in ancillary charges for those who qualified for Extensive Services were observed across all five Rehabilitation categories.

These cost differences suggest a potential type of refinement for residents who qualify for both Extensive Services and Rehabilitation. Such a refinement could be implemented by making fairly minor changes to the structure of RUG-III. For example, a new category could be added for these residents. If the structure of these categories were identical to that of the current RUG-III Rehabilitation categories, the resulting casemix system would have 14 additional Extensive Services

and Rehabilitation Groups, which would use the same Rehabilitation categories and ADL splits as the current Rehabilitation categories.[3] We refer to this refinement as the RUG-III+ model.

There are fairness and consistency-related reasons for considering changes to the casemix system for residents who qualify for both Extensive Services and Rehabilitation. Given that the payment for other types of residents is tied to the amount of therapy received, the inconsistency of having payment for Extensive Services invariant to the amount of therapy received (which would be the case if the Extensive Services payment rate were adjusted to reflect the non-therapy ancillary charges observed for the category) raises concerns about the appropriateness of treating Extensive Services residents differently than other types of residents. The structure of the RUG-III Rehabilitation categories reflects a decision by HCFA to encourage SNFs to provide therapy (by having the payment rate tied to the amount of therapy received) and it would be inconsistent to treat residents in the Extensive Services category differently.

3 It should be noted that RUG-III+ could also be implemented as a new terminal split within the existing Rehabilitation categories based on whether the resident also qualified for Extensive Services. Statistically, the two systems are identical.

Table 5.3

Comparison of Total Costs for Those in Rehabilitation Categories Based on Whether Resident Also Qualifies for Extensive Services

RUG-III category	Qualifies for Extensive Services		Does not qualify for Extensive Services	
	N	Mean ancillary charges (std. dev.)	N	Mean ancillary charges (std. dev.)
All Rehabilitation categories	2,926	$119.13 (172.65)	26,699	36.74 (67.82)
Ultra-high Rehabilitation	496	175.87 (207.21)	4,825	44.14 (76.74)
Very-high Rehabilitation	446	124.06 (177.40)	4,675	42.16 (77.76)
High Rehabilitation	379	115.40 (175.08)	4,169	33.76 (66.41)
Medium Rehabilitation	1,530	101.27 (153.86)	11,993	33.71 (61.51)
Low Rehabilitation	75	97.59 (164.73)	1,037	24.82 (40.36)

Notes: N=29, 625 (Based on test sample residents in a RUG-III Rehabilitation category only)

Data Source: Medicare MDS and SNF Claims Data 1995-1997

5.4 MDS Items Associated With Differences in Ancillary Charges

Given the large within-category variance in ancillary charges, the next phase of our analyses was dedicated to identifying MDS items associated with differences in prescription drug, respiratory therapy, or other non-therapy ancillary charges. We first identified MDS items that were associated with either higher drug charges, respiratory therapy charges, or higher other ancillary charges. These items were identified based on a broad and systematic search of MDS items (see Appendix C).

Our work suggested that refinements based on individual MDS items produced only very small improvements in statistical performance, and instead focused our efforts on refinements based on combinations of MDS items (i.e., index models). Refinement efforts also focused on predicting total ancillary charges. Some of the variables that were associated with differences in either drug, respiratory therapy, or other ancillary charges were not associated with differences in *total* ancillary

charges, and these variables were dropped from the index model. The MDS items that were associated with differences in total ancillary charges are reported in Table 5.4. This table reports both means and regression coefficients associated with each variable. The means indicate how prevalent each treatment or condition was in the data, for example 19 percent of the sample were receiving oxygen therapy. The regression coefficients measure each variable's contribution to total ancillary charges, holding other variables constant. For example, receipt of oxygen therapy was associated with $21.22 in additional charges (after discounting).

Based on this search, we identified a subset of items from the MDS that had a significant relationship with ancillary charges. These items include a variety of measures of resident acuity and treatments received, including functional status (bedfast), nutritional status (parenteral/IV feeding, tube feeding), disease diagnoses (COPD, terminal condition, pneumonia, respiratory infection), health conditions (use of indwelling catheter, shortness of breath), skin conditions and treatments (Stage 4 pressure ulcers, surgical wound/ulcer care, application of dressing with/without topical medication), and special treatments and procedures (IV medications, tracheostomy, suctioning).

Table 5.4

MDS Items Associated With Higher Total Ancillary Charges

MDS Item	Mean	Regression Coefficient (std. error)	Implications
Parenteral /IV	0.025	84.61 (2.41)	IV Hydration presents some opportunity for manipulation; however, one safeguard is to pair with amount of parenteral nutrition provided via IV so that costs captured are those related to identified nutritional need.
Suctioning	0.022	71.70 (2.86)	Labor and supply-intensive item, not likely to be manipulated.
Tracheostomy Care	0.015	48.44 (3.29)	No apparent negative incentive.
IV Medication	0.118	42.68 (1.18)	Currently included in Extensive Care; potentially provides incentive to administer medication generally available in oral or injectable form as an IV.
Oxygen	0.190	21.22 (1.45)	Without linking oxygen use to a diagnosis/condition and/or symptoms indicative of need could lead to inappropriate and over use of oxygen.
COPD	0.214	20.98 (0.96)	Due to prevalence of this diagnosis, potential for manipulation as a stand-alone variable. Needs to be linked to treatment (oxygen) and/or symptoms of acuity (SOB).
Terminal Condition	0.021	10.61 (2.57)	Eliminated from RUG-III.
Pneumonia	0.104	15.87 (1.27)	Less likely to be manipulated; could be linked to oxygen use, fever and/or SOB if gaming a concern.
Tube Feeding	0.095	15.57 (1.36)	Item currently included in Clinically Complex. Needs to be linked to percentage of calories and fluids received via the tube to avoid negative incentive affect.
Shortness of breath	0.176	15.90 (1.15)	Subjective with probable high prevalence. Highly gameable.
Bedfast	0.138	11.19 (1.19)	Negative incentive to quality of life/quality of care.
Number of Stage 4 Pressure Ulcers	0.041	2.15 (0.54)	Item included in Special Care, and is a Quality Indicator. Unlikely to provide negative incentives.
Respiratory Infection	0.075	12.32 (1.46)	Not likely to be manipulated.

Table 5.4

MDS Items Associated With Higher Total Ancillary Charges

MDS Item	Mean	Regression Coefficient (std. error)	Implications
Application of dressing with/without topical medication	0.056	8.68 (1.63)	Includes wide variety of dressings/bandages, not all of which are high cost. Should be paired with indication of wound requiring dressing. Currently included in Special Care.
Surgical Wound/Ulcer Care	0.282	4.08 (0.85)	Item is included in Special Care.
Indwelling Catheter	0.177	10.71 (1.04)	Creates facility incentives for negative clinical outcomes. Would only consider the variable if linked to certain diagnoses/conditions.

Notes: N = 52,328 (Test sample only)

Data Sources: Medicare MDS and SNF Claims Data 1995-1997

5.5 Clinical Review and Modification

The items listed in Table 5.4 were identified using statistical techniques, and some of the items associated with significant cost differences may be inappropriate for use in the casemix system due to clinical or incentive-related concerns. Based on input from a Clinical Workgroup assembled for this project and the clinical expertise of project staff, we modified the list of items to include in the index model. Some items, despite their ability to identify high cost residents, were rejected outright due to potential negative incentive effects. Others were found acceptable with modification, and the remainder were recommended as is for inclusion in a potential model. A further description of this review follows.

MDS Items Considered Unacceptable

The MDS item indicating that a resident is "bedfast" is not considered appropriate for inclusion in a casemix classification system. Though the index model shows that sampled residents who are bedfast have $12.00 higher *per diem* non-therapy ancillary charges than those residents who are not bedfast, this item could be conceived as encouraging facilities to allow residents to spend excessive time in bed or in their room and thus have a negative impact on facility staffing, resident quality of life and quality of care. Similarly, indwelling urinary catheters are best excluded from the classification system in order to avoid their inappropriate use. Only by linking to medication profiles and diagnoses would these items be considered appropriate for inclusion in any casemix classification system. The Terminal Prognosis variable was removed from RUG-III prior to national PPS implementation, due to it's lack of predictive power and incompatibility with the Medicare SNF benefit.

MDS Items Found to be Acceptable if Linked to Other MDS Items

Several of the most powerful MDS variables identified by the index model indicate either treatment or symptoms/conditions which are highly prevalent in the nursing home population. Such variables when considered independently are of some concern, as they have the potential to create perverse incentives. In order to make use of these variables while maintaining some assurance that the modified classification system will not promote negative clinical outcomes, the following combinations of variables were proposed by the clinical team.

Parenteral/IV – This item includes intermittent fluid administration for hydration and its inclusion could lead to unnecessary invasive practices. If utilized in the classification system, the item should focus on parenteral nutritional fluids administered either via central or peripheral lines. The item could be refined by linking it to the percentage of calories received via parenteral IV, without the presence of a feeding tube.

Oxygen – Because oxygen administered for any brief period in the last 14 days generates coding on this item, there is concern is that – absent a link to an acute condition – there could be inappropriate overuse of oxygen. We propose linking oxygen use to the following diagnoses and symptoms:

Diagnosis	Symptoms
Respiratory Infection and shortness of breath (SOB)	Fever
Pneumonia and SOB	Fever
COPD	SOB
CHF and Inability to lie flat	SOB
CAD and Inability to lie flat	SOB
Terminal Illness	SOB
Cancer and Terminal Illness	SOB

Shortness of Breath – Because of the subjectiveness and wide scope of this item, we believe it presents high opportunity for manipulation. To narrow the focus of this item to only those residents with significantly acute clinical conditions, the variable should be further evaluated by linking to: COPD; Pneumonia with Fever; Respiratory Infection with Fever; Terminal Illness; Cancer with Terminal Illness; or CHF with Inability to lie flat.

COPD/Emphysema – Because of the widespread prevalence of this condition, the item should focus on acute exacerbations. Combining these diagnoses with MDS items that reflect treatment of an acute condition would minimize potential gaming. A link to oxygen use and SOB would further refine this item.

Feeding tube – This item is included in Clinically Complex and by RUG group definition is linked to percentage of calories and fluids received via the tube (which we would support in further RUG refinement).

Dressing application with/without topical medications – This item may include a wide array of dressing types, not all of which represent the need for unusually high amounts of supplies or

medications. Linking to specific clinical conditions would better define this item in terms of non-therapy ancillary charges. The suggestion was made to evaluate this item in combination with items indicating the need for treatment with dressings, i.e. presence of ulcers or other wounds or lesions.

Based upon suggestions from our clinical team (and from the earlier Clinical Workgroup), several MDS variables were examined to determine their interactions. Specifically, the index models were re-run with four re-defined variables:

- Parenteral IV with the majority of caloric intake (greater than 75 percent) being administered via parenteral IV;
- Oxygen administration and either pneumonia or respiratory infection with fever; or oxygen with pneumonia, respiratory infection, COPD, CHF, or CAD – all with shortness of breath;
- Feeding tube with the majority of caloric intake (greater than 75 percent) being administered via feeding tube; and
- Application of dressing with presence of either ulcers or other skin lesions/wounds.

MDS Items Acceptable "As-Is"

In general, the following MDS items by themselves do not appear to provide negative care incentives nor would they be considered "gameable": IV Medications, Suctioning, Respiratory Infection, Pneumonia, Tracheostomy Care, and Stage 4 Pressure Ulcers.

Suggestions for MDS Form Modification

In order to better capture the types of treatments and corresponding conditions we suspect contribute to the higher non-therapy ancillary charges associated with the items described above, we would suggest several modifications to the MDS assessment form in future revisions. These proposed modifications were generated through discussion with the Clinical Workgroup and through internal clinical review of potential RUG-III refinements.

- A clear distinction between types of IV access (e.g., central v. peripheral line); reason for IV (e.g., antibiotics, medication including IV push medication, hydration, nutrition (TPN or PPN);
- Re-evaluation of the time frames for measurement of diagnoses and infections and occurrence of symptoms. For example, when attempting to link treatment to clinical conditions and symptoms, we found that MDS assessments could be incorporating data from three different time periods, thereby raising the possibility that we may not be measuring the same condition across the various time periods, but rather three separate conditions. Also regarding time frames, all the items in Special Treatments (P1a Special Care) refer to treatments that were provided within the past 14 days, not necessarily treatments provided at the nursing facility. A distinction between treatments provided prior to admission from those provided at the

facility would provide a more accurate description of care/cost associated with the nursing facility.

- A clearer definition of pressure relieving bed and chair, as this item is believed to be over-utilized and is one of several skin treatments that, when present, allow a resident to meet the requirement for a Special Care category.

- Finally, the Clinical Workgroup noted that inclusion of information on the location of skin ulcers (i.e., trunk vs. extremity) would enable differentiation between appropriate/necessary means of treatment vs. unnecessary and/or costly methods.

Results of Refined Index Model

The items in the refined index model performed reasonably well, accounting for 18 percent of the variance in ancillary charges for the three states (Mississippi, South Dakota, and Texas) for which the caloric intake MDS items was available (This item was available only for MDS2.0 data). The addition of clinical conditions to oxygen administration and of calories to parenteral/IV did not substantially undermine the predictive power of the model. A six-state version of the index model that did not use the interaction between caloric intake and parenteral/IV or tube feeding predicted 18 percent of the variance in ancillary charges. Regression coefficients for both the three- and six-state versions of the model are reported in Table 5.5.

Only 32 percent of residents who received parenteral/IV feeding met the caloric intake threshold recommended by the Clinical Workgroup. Average ancillary charges for these residents were $372, compared to $216 for the 46 percent of those who were reported as receiving parenteral/IV feeding but zero percent of calories from parenteral/tube feedings. Nearly 84 percent of those who used a feeding tube met the caloric intake threshold.

Regression coefficients for the modified index model are reported in Table 5.5. Two sets of estimates are presented– one based on the three states for which the parenteral/IV and tube feeding interaction variables could be created and a second which included all states but did not include these interactions. Because of serious concerns about how representative the cost estimates from the three state model were, refinements were evaluated using the six-state model. For example, the three state sample included only 8 non-Texas residents with the parenteral/IV feeding- caloric intake variable (there were 98 Texas residents with this item present.)

Table 5.5

MDS Items Associated with Differences in Ancillary Charges– Refined Variable List Following Clinical Input

	6-state version of model†		3-state version of model†	
MDS Item	**Mean**	**Regression coefficient (std. error)**	**Mean**	**Regression coefficient (std. error)**
Intercept	–	24.18 (0.44)	--	32.69 (1.07)
Parenteral /IV with > 76% total calories†	0.027	79.25 (2.13)	0.006	225.22 (10.81)
Tracheostomy Care	0.06	52.53 (2.87)	0.006	154.61 (11.11)
Suctioning	0.022	79.32 (2.55)	0.021	107.05 (6.22)
IV Medication	0.117	44.83 (1.07)	0.160	58.15 (2.27)
Oxygen and either pneumonia or resp. inf. with fever, or pneumonia or resp. infection, COPD, CHF, CAD with SOB	0.099	44.06 (1.15)	0.098	61.95 (2.86)
Pneumonia	0.107	17.54 (1.15)	0.110	23.99 (2.73)
Tube Feeding with > 76% total calories†	0.101	18.13 (1.16)	0.067	35.25 (3.39)
Respiratory Infection	0.079	14.80 (1.31)	0.072	24.61 (3.28)
Application of dressing with/without topical medication and presence of ulcers or other skin lesions/wounds	0.052	11.12 (1.55)	0.061	7.14 (3.60)
Skin Wound/ Ulcer Care	0.028	5.66 (0.77)	0.263	10.66 (1.96)
Stage 4 Pressure Ulcer	0.055	0.80 (0.23)	0.063	11.44 (2.08)

Notes: N: 17,788 for three state model, 52,328 for six state model (Based on test sample only).

†: Due the unavailability of the MDS 2.0 item that reports the percentage of total calories from parenteral or tube feedings for assessments from 1995 as well as assessments from Kansas, Maine, and Ohio, we also estimated a three-state version of the model that did not use the interaction variables for parenteral/IV and the percentage of calories and tube feeding and the percentage of calories. Instead, the six state version of the model used only indicators for whether the residents received parenteral/IV feeding or tube feeding.

Data Source: Medicare MDS and SNF Claims Data 1995-1997, excluding ME, OH, SD

5.6 Description of Potential RUG-III Refinements

Given the high ancillary charges for residents who qualified for both Rehabilitation and Extensive Services, we evaluated refinements based on the creation of a new category for residents who qualify for both Extensive Services and a Rehabilitation category. This refinement produces some improvement in the ability of the casemix system to account for variance in ancillary charges using "internal changes" that are based on the existing RUG-III categories. In addition, the statistical performance of the casemix system is improved through the use of index used either to create new groups for selected RUG-III categories or as a separate "ancillary add-on." The exact type of index model to use is largely a policy decision based on trade-offs between improved statistical performance and increased casemix system complexity. There are a number of ways that index model-based refinements can be implemented, several of which are described below.

Development of index model-based refinements

All of the index model-based refinements incorporate new categories for residents who qualify for both Extensive Services and a Rehabilitation category. There are a number of ways that index model-based refinements can be implemented:

- The models can be based on an unweighted count of the number of index model variables present or on a weighted index that uses the index model variables to estimate predicted ancillary charges.

- The index models can differ with respect to the RUG-III categories to which the model is applied. The two best options seemed to be to apply the index model only to Extensive Services residents (including Rehabilitation residents who qualify for Extensive Services) or to apply it to all residents in Clinically Complex or above.

- The index models can differ with respect to the number of index groups that are used. For the weighted index models described below, residents were put into one of six groups based on their predicted ancillary charges. For the unweighted model, four groups were used. The number of index groups can be changed, but there are costs, in terms of decreased statistical performance, that result from decreasing the number of groups.

- The index models can also vary based on the thresholds used to define groups. For the weighted index model, residents were classified based on whether their predicted costs were at the 50th percentile or below (group 1), the 51st - 75th percentile (group 2), the 76th-90th percentile (group 3), the 90th - 95th percentile (group 4), the 96th - 98th percentile (group 5) and the 99th percentile (group 6). For the unweighted index model, groups were defined based on whether zero, 1-2, 3-5, or 6 or more index model variables applied to the resident.

- For the index model alternatives described below, the same index was used across all RUG-III categories, but it is also possible to apply category-specific index models.

We tested a number of potential refinements, but focus on the models described below. Since we considered several types of refinements in depth, we selected only the most powerful alternative from each type for presentation here. These were:

- *Model RUG-III+*: This is the RUG-III model with new categories for residents who qualify for Extensive Services and one of the RUG-III Rehabilitation categories, as described above. The resulting casemix system would have 14 additional Extensive Services and Rehabilitation Groups, which would use the same Rehabilitation categories and ADL splits as the current Rehabilitation categories.[4]

- *Model WIM 1:* Applies the weighted index model to Extensive Services residents (including residents in the new Extensive Services and Rehabilitation categories). A disproportionate share of high cost residents qualify for Extensive Services, so it may make sense to apply the index model only to those residents. Using this refinement, the casemix system would have up to 143 groups if the index model were incorporated within RUG-III as new terminal splits. Alternatively, the system could be implemented as a six-group ancillary add-on system.

- *Model WIM 2:* Applies the weighted index model to Extensive Services residents (including residents in the Extensive Services and Rehabilitation category), and to Rehabilitation, Special Care, and Clinically Complex residents. In this model, there would be up to 258 groups if it were implemented as new terminal splits rather than as a 6-group ancillary add-on system.

- *Model UWIM:* Applies the unweighted index model to Extensive Services residents (including residents in the Extensive Services and Rehabilitation category), and to Rehabilitation, Special Care, and Clinically Complex residents. In this model, there would be up to 178 groups if it were implemented as new terminal splits rather than as a 4-group ancillary add-on system.

(Note that for comparison purposes, we also evaluated the baseline RUG-III model, with changes only in the CMI.)

Although the proposed index model variables include interactions for parenteral/IV feeding and feeding tubes based on the proportion of calories that the resident obtained through parenteral feeding or feeding tubes, *this interaction was not included in analysis of the statistical performance of potential refinements.* The MDS item that reports proportion of calories from IV/feeding tube was available only for residents in Mississippi, South Dakota, and Texas. (Note that New York was excluded from for the reasons described in Chapter 4.) We believed that it is important to develop and test potential refinements on a broader sample. Our analysis suggested that the statistical performance of the two models was comparable. *Although the interaction variables were not included in the potential refinements that we evaluated, we recommend that they be included into the refined casemix system.* The interaction variables are important for reducing the extent to which providers might inappropriately use parenteral/IV feeding or feeding tubes solely for the purpose of receiving a higher

4 It should be noted that RUG-III+ could also be implemented as a new terminal split within the existing Rehabilitation categories based on whether the resident also qualified for Extensive Services. Statistically, the two systems are identical.

payment. Further research using a larger sample of MDS2.0 assessments is needed, however, to calculate reliable estimates of the cost differences associated with these interaction items.

Model performance

RUG-III: Changes to the CMI alone (i.e., changes to the payment rates associated with different groups but no changes to the casemix system) will reduce the proportion of individuals for whom costs are greater than payment, but will not affect the proportion of variance in costs captured by the casemix system. RUG-III accounted for about 4 percent of the variance in ancillary charges and 10 percent of the variance in total costs (Table 5.6). Predicted costs for the 44 RUG-III groups ranged from $111 to $267 (for SE3). Using RUG-III, 22 percent of residents in the top 10 percent in terms of actual ancillary charges were also in the top 10 percent in terms of predicted ancillary charges.

RUG-III+: Adding the new Extensive Services and Rehabilitation categories (model RUG-III+) resulted in small improvements in statistical performance. The validation sample R-squared increased to 8 percent for ancillary charges, an increase of about 4 percent relative to RUG-III, and to 14 percent for total costs. Predicted costs for the 58 groups in the RUG-III+ model ranged from $116 (excluding a behavior problem group which had very few residents and an imprecise predicted cost estimate) to $385 (for one of the new Extensive Services and Ultra-high Rehabilitation groups). RUG-III+ was statistically less powerful than the other refinements we examined, all of which incorporated RUG-III+ and index-index model-based refinements. The RUG-III+ refinement, however, adds less complexity than the index-model based refinements.

WIM 1: Application of WIM 1 resulted in some statistical improvement relative to RUG-III+. For the validation sample, the model accounted for 12 percent of the variance in ancillary charges and 18 percent of the variance in total costs. For both WIM1 and RUG-III+, 29 percent of residents in the top 10 percent of ancillary charges were also in the top 10 percent in terms of predicted costs.

Under WIM1, Extensive Services residents (including those in the new Extensive Services and Rehabilitation categories) would receive an ancillary "add-on" based on the index model variables applicable to the resident. A six group ancillary index was used. There would be no additional ancillary "add-on" for residents whose predicted costs are below the 50th percentile in terms of predicted ancillary charges, $17 for those between the 51st and 74th percentile, $34 for those in the 75th - 89th percentile, $56 for those in the 90th - 94th percentile, $106 for those in the 95th - 98th percentile, and $225 for those in the top 1 percent. The cutoff points for the index models are weighted towards high cost residents since those residents account for much of the variance in ancillary charges. Because of the emphasis on very high cost residents, the top payment rate (based on predicted costs) is $520 (for residents in one of the Extensive Services and Ultra-high Rehabilitation groups who are also in the highest index group (i.e., in the top 1 percent in terms of predicted ancillary charges). This was more than $250 higher than the top predicted cost under RUG-III, showing the improved potential of the model to capture costs for residents with high resource needs.

WIM 2: The WIM2 model accounted for 22 percent of the variance in total costs and 16 percent of the variance in ancillary charges. The range of payments was similar to that of WIM1. Using WIM2, 34 percent of residents in the top 10 percent in terms of actual ancillary charges were also in the top 10 percent in terms of predicted ancillary charges, a substantial improvement relative to RUG-III.

Model WIM2 was the most statistically powerful refinement that we examined. Because the model applies a six-group ancillary index to 40 RUG-III+ groups (14 Rehabilitation/Extensive Services groups, 3 Extensive Services groups, 14 rehabilitation groups, 3 Special Care and 6 Clinically Complex groups), it results in a large number of groups if it is implemented as part of an integrated classification system. Alternatively, as with the other index model-based refinements, WIM2 could be thought of as a six group ancillary add-on which works alongside RUG-III to determine total payment.

UWIM: This model is the unweighted counterpart to WIM2 and is based on a count of the number of index model items present rather than the cost difference associated with each item. While this model performed better than the RUG-III and RUG-III+ models, it did not perform as well as WIM2. The model would be favored over WIM2 if there were concerns about the complexity of the weighted index models or other concerns about the weighted models, perhaps because the cost data covered a period before implementation of PPS.

UWIM accounted for 12 percent of the validation sample variance in ancillary charges and 19 percent of the variance in total costs. The range of payments for UWIM was quite similar to that of the weighted index models. The sensitivity and specificity of the model were slightly less than for WIM2.

Using UWIM, residents were split into four groups based on the *number* of index model variables applicable to each resident. The splits used were 0 (45 percent of test sample observations), 1-2 (45 percent), 3-5 (9 percent) and 6 or more (0.4 percent). Residents with no index model items present would receive no additional ancillary payment, while those with 1-2 items present would receive $19 (based on predicted ancillary charges), those with 3-5 items would receive $68 and those with six or more would receive $209.

Table 5.6

Statistical Performance of Potential RUG-III Refinements— Model Description

Model description	Number of groups	R-squared Validation sample (Test sample)		Min/ Max▲	Specificity and sensitivity analyses Validation sample	
		Ancillary charges	Total costs		Specificity*	Sensitivity◇
RUG-III (CMI changes only)	44	4.1% (3.5%)	10.0% (9.3%)	117/267	91.4%	22.2%
RUGIII+ *RUG-III with new category "Extensive Services and Rehabilitation"*	58	8.0% (7.0%)	14.3% (12.3%)	116/385	91.6%	29.4%
WIM 1 *Weighted index model applied to Extensive Services residents (includes new category "Extensive Services and Rehabilitation")*	58 plus a six-group ancillary add-on system (Up to 143 if incorporated as new terminal splits)	12.0% (9.4%)	18.1% (15.7%)	118/520	92.1%	28.7%
WIM 2 *Weighted index model applied to Extensive Services residents (includes new category "Extensive Services and Rehabilitation") and to Rehabilitation, Special Care, and Clinically Complex residents*	58 plus a six-group ancillary add-on system (Up to 258 if incorporated as new terminal splits)	15.5% (13.6%)	21.9% (20.2%)	116/520	92.6%	33.5%
UWIM *Unweighted index model applied to Extensive Services residents (includes new category "Extensive Services and Rehabilitation") and to Rehabilitation, Special Care, and Clinically Complex residents*	58 plus a four-group ancillary add-on system (Up to 178 if incorporated as new terminal splits)	12.3% (11.1%)	19.0% (18.0%)	116/517	92.3%	30.6%

Notes: ▲: Predicted total costs for the lowest (assumed to be PA1) and highest reimbursed groups in the refined casemix system (uses simulated staff time costs derived from HCFA's Staff Time Measurement Studies.)

★: Specificity is measured as the proportion of residents who are not in the top 10 percent of predicted ancillary charges and also not in the top 10 percent in terms of actual ancillary charges.

◇: Sensitivity is measured as the proportion of residents in the top 10 percent in terms of both predicted and actual ancillary charges.

Data sources: Medicare claims, Minimum Data Set 1995-1997

Limitations to the Analysis

The potential refinements described in Table 5.6 result in some improvement over the RUG-III system, but the results are only moderately successful. The most powerful model that we were able to design achieved rather modest predictive power, explaining less than 14 percent of the variance in ancillary charges and less than 20 percent of the variance in total costs. While we are unable to do better, we nevertheless suggest implementation of one of the recommended models. The potential refinements do lead to improved statistical performance and allow for higher payment rates for residents with characteristics associated with high ancillary charges, addressing some significant portion of the industry's valid concerns.

It is informative to understand the reasons for the relatively modest results that were observed. We suggest that they are the result of several phenomena. Two have the result of increasing the "noise", or measurement error, in either the dependent variable (costs) or the independent variables (the MDS items), reducing the predictive power of RUG-III and potential refinements.

- *Measurement error in the measure of ancillary charges* Deriving accurate measures of drug costs from Section U or measures of ancillary charges from Medicare claims data have intrinsic problems (see discussion in Chapter 2 regarding limitations in ancillary charge measures based on SNF claims and reasons for differences between drug cost measures based on SNF claims and Section U.)

- *Measurement error in resident characteristics* By using a population-based sample, we can expect that some MDS assessments will be well done, but a substantial portion may have been inaccurately completed, especially for items that do not affect payment rates. In the earlier Abt study, the statistical performance of potential refinements similar to those described above was somewhat higher. For example, a model similar to WIM2 achieved a variance explanation of 24 percent (validation sample) in the earlier Abt study, compared to 19 percent in the current study. This may partly be due to lower quality MDS data. The MDS assessments used in the earlier study were collected as part of HCFA's Staff Time Measurement Study and were likely to be more accurately completed than the MDS assessments used for this study. This suggests that if particular items are identified in the casemix system as related to reimbursement, the accuracy of these items will increase and the eventual explanation of actual cost can be expected to increase concurrently.

Our models cannot explain that part of the variation in ancillary charges that is due to errors in ancillary charges derived from Medicare claims and measurement error in the MDS assessments also reduces statistical performance. We thus do not know what part of the *true* variation in ancillary charges we are explaining. In addition to reductions in statistical performance resulting from measurement error and noisy data, we know that there is considerable variation in ancillary charges across states (See Appendix D) and facilities (e.g., hospital-based, for-profit/non-profit, part of chain etc.). These items are associated with differences in ancillary charges, but are not appropriate for inclusion in index models or other types of casemix refinements. The portion of ancillary charge variance that is due to state and facility effects, as well as other variables that could not be included in

our models is thus beyond the reach of our models, reducing statistical performance. We are anxious to expand upon our models to control for some of these effects, thereby giving a more accurate picture of the true between-resident variation in costs.

References

Abt Associates Inc. Refining Resource Utilization Groups (RUG-III) for a National Skilled Nursing Facility System - Final Report to HCFA; March 1999.

Austin Data Management. PC-Group. Austin, TX; Austin Data Management, 1992.

Cameron JM. Casemix and Resource Use in Long-Term Care. Medical Care 1985; 23:296-309.

Cornelius E, Feldman J, Marsteller JA, Liu K. Creating a MEDPAR Analog to the RUG-III Classification System. Health Care Financing Review 1994;16(2):101-126.

Fetter RB, et al. The New ICD-9-CM Diagnosis-Related Groups Classification Scheme, User Manual, Vol. 1, Yale University, New Haven, CT., 1981.

Fries BE and Cooney LM. Resource Utilization Groups: A Patient Classification System for Long-Term Care. Medical Care 1985; 23:110-122.

Fries BE, Schneider D, Foley WJ, Gavazzi M, Burke R, Cornelius E. "Refining a Casemix Measure for Nursing Homes: Resource Utilization Groups (RUG-III)" Medical Care 32(7):668-685 (July) 1994.

Gambassi G, Landi F, Peng L, Brostrup-Jensen C, Calore K, Hiris J, Lipsitz L, Mor V, Bernabei R. Validity of diagnostic and drug data in standardized nursing home resident assessments. Medical Care 1998;36(2): 167-179.

Katz S, Form AB, Moskowitz RW, et al. Studies of Illness in the Aged: The Index of ADL: A Standardized Measure of Biological and Psychological Function. Journal of the American Medical Association 1963; 195:914-919.

Minnesota Department of Health. Facility Manual for Completing Casemix Requests for Classification. Minneapolis, Minnesota, 1986.

Morgan JN and Sonquist JA. Problems in the Analysis of Survey Data and a Proposal. J. American Statistical Association 1963; 58:415-434.

Morris JN, Fries BE, Mehr DR, et al. MDS Cognitive Performance Scale. Journal of Gerontology 1994; 29(4):M174-M182.

Morris JN, Sherwood S, May MI, et al. An Innovative Approach to Nursing Home Level-of-Care Assignments. Health Services Research 1987; 25:9-19.

Morrison Informatics Inc. and University of Colorado Health Sciences Center. A Prospective Study of New Casemix Indices for Subacute Care - Final Report; June 1999.

Schneider DP, Fries BE, Foley WJ, et al. Casemix for Nursing Home Payment: Resource Utilization Groups, Version II. Health Care Financing Review 1988; (Supp):39-52.

Sherwood, J.D. Telephone Conversation. December 6, 1999.

Williams BC, Fries BE, Foley WJ, Schneider D, Gavazzi M. "Activities of Daily Living and Costs in Nursing Homes" Health Care Financing Review 15(4):117-135 (Summer) 1994.

U.S. Department of Health and Human Services Office of Inspector General. Early Effects of the Prospective Payment System on Access to Skilled Nursing Facilities. August 1999.

U.S. General Accounting Office. Skilled Nursing Facilities: Medicare Payments Need to Better Account for Nontherapy Ancillary Cost Variation. September 1999.