

U.S. Department of Health and Human Services
Centers for Medicare & Medicaid Services (CMS)

Report to Congress:
Post Acute Care Payment Reform
Demonstration (PAC-PRD)

January 2012

Table of Contents

INTRODUCTION	1
SECTION 1 UNDERLYING ISSUES OF THE PAC-PRD INITIATING LEGISLATION	2
1.1. Ability to Compare Patients across Settings	3
1.2 Inconsistencies in Case-Mix Systems and Unintended Incentives	5
SECTION 2 DATA COLLECTION IN THE DEMONSTRATION	6
2.1 General Data Collection Approach	6
2.2 Development of a Standardized Patient Assessment Tool	8
SECTION 3 RESOURCE INTENSITY ANALYSIS RESULTS	12
SECTION 4 OUTCOMES ANALYSIS RESULTS	28
SECTION 5 REVIEW OF RESULTS AND RECOMMENDATIONS	34
5.1 Key Results in the Resource Intensity Analysis	35
5.2 Key Results in Outcomes Analysis	36
5.3 CARE tool after the Demonstration	37
APPENDIX 1: SEC. 5008. POST-ACUTE CARE PAYMENT REFORM DEMONSTRATION PROGRAM. DEFICIT REDUCTION ACT OF 2005	39

List of Tables

Table 1: Count of participating providers, by provider type, by market area	8
Table 2: CARE assessment counts by assessment form, by provider type	9
Table 3: Number of cleaned CRU patients, patient-days, and days per patient by setting	13
Table 4: Ratio of Predicted to Actual Resource Intensity Index by Setting, for Therapy and Routine Stay Level RII Alternative Models	24

INTRODUCTION

In section 5008 of the Deficit Reduction Act of 2005 (hereafter referred to as the “DRA of 2005”), Public Law 109-171, Congress required the Post-Acute Care Payment Reform Demonstration (PAC-PRD) and directed CMS to deliver a report on the results of the program and recommendations for such legislation and administrative action as the Secretary determined to be appropriate. The key components of the demonstration included:

- Participating facilities included the four types of Post Acute Care (PAC) providers covered under Medicare Part A: Long Term Care Hospitals (LTCHs), Inpatient Rehabilitation Facilities (IRFs), Skilled Nursing Facilities (SNFs), and Home Health Agencies (HHAs). In addition, information was collected at discharge in subsection (d) hospitals (as defined in section 1886(d)(1)(B) of the Social Security Act (42 U.S.C. 395ww(d)(1)(B))). This provider type will be referred to as acute care hospitals (ACH) for the purposes of this report;
- Data collection included a standardized patient assessment instrument for discharge from acute care hospitals and at admission and discharge from post acute care sites. The standardized patient assessment instrument provided information on clinical and other patient factors associated with costs and resource use, outcomes, discharge placement and care transitions;
- Data collection in PAC-PRD included measures of patient specific costs and resource expenditures; and,
- The analyses focused on increasing understanding of factors that predict costs and outcomes in each of the different post acute care sites.

Work on the PAC-PRD was divided into three major components. Component One developed a single comprehensive patient assessment instrument that could be used in acute care hospitals and the four PAC settings of interest. The work resulted in the creation of the CARE (Continuity Assessment Record and Evaluation) tool. Component Two developed a secure internet based software application for collecting CARE assessment information from the

participating providers. Finally, Component Three collected data and performed the associated analyses of the demonstration.

This report summarizes the PAC-PRD's major activities, results, and recommendations. Section 1 covers underlying issues leading to the need for this project. Section 2 briefly describes the development of a common patient assessment tool and the data collection methods. Section 3 presents selected results from the resource intensity analysis. Section 4 covers outcomes results. Finally, Section 5 reviews the key results from the project and presents next steps. A more detailed Contractor Report will be posted on the CMS website.

SECTION 1
UNDERLYING ISSUES OF THE PAC-PRD INITIATING LEGISLATION

Currently, Medicare beneficiaries may be discharged from an acute care hospital to one of four PAC settings for additional nursing or therapy treatments, including LTCHs, IRFs, SNFs, or HHAs. In general, the four PAC settings are assumed to differ in the type and intensity of services provided and, for the most part, provide a "continuum of care." However, many conditions may be treated in multiple PAC settings, therefore, there is some degree of overlap in the populations treated. This prompted a study using a standardized patient assessment instrument to examine the amount of overlap in patient acuity, resource use, and outcomes across PAC sites.

The various PAC payment systems differ considerably in how they incorporate and measure the proportion of costs that vary with patient characteristics. These varying approaches have led to different classification of and payment for patients even though the clinical characteristics of the patients and the services delivered may be very similar across the settings. While fixed costs understandably vary among hospitals, facilities, and agencies, the difference in

case-mix adjusted payments among settings can lead to inefficient patterns of care within a trajectory of care¹. Evaluating the populations served in each of the four settings and the resources needed to care for these patients using the data collected in this demonstration is an important step in assessing whether payment approaches that have greater consistency between sites of care show promise.

1.1. Ability to Compare Patients across Settings

An important issue addressed by this demonstration is the potential variation in payment levels across provider settings for the same type of patient. In order to examine variation in payment levels, one must be able to assess and risk-adjust patient severity and needs in a consistent manner between settings.

The Medicare program currently mandates that IRFs, SNFs, and HHAs submit assessment data on the beneficiary's medical, functional, and cognitive status. These mandated site-specific patient assessment tools are the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI), the Minimum Data Set (MDS), and the Outcome and Assessment Information Set (OASIS), respectively. These instruments are used to measure patient severity at admission and during different times over the course of the patient's treatment. The information collected through these assessments is used by CMS to calculate payment groups, generate quality measures, and monitor regulatory compliance. Additionally, many states use data from these assessments for Medicaid payment and quality monitoring.

¹ A trajectory of care is an informal term used to refer to the care a patient receives in order to treat a spell of illness associated with a hospitalization. A trajectory may include one or more settings. The term "trajectory" is used rather than "episode" in order to distinguish it from the home health Prospective Payment System (PPS) use of the term episode, and other research projects that use the term episode in a different manner related to the treatment of chronic conditions. The use of the term "episode of care" in section 5008 of the DRA of 2005 is interpreted to be consistent with this use of "trajectory of care."

Although the three mandated assessments measure similar concepts, they use different clinical items, different assessment timeframes and disparate measurement scales to assess health, physical function, and cognitive status. Acute care hospitals and LTCHs collect similar information at intake but are not currently required to submit these data to CMS. Instead, payments to acute care hospitals and LTCHs are based on claims information which identifies the precipitating acute event and associated procedures.

Similar patients can be treated in more than one provider setting. The ability to consistently measure patient acuity, resource use, and outcomes across settings will help to guide appropriate policies for these patient populations. Changing practice patterns have promoted the unbundling of care historically delivered in an ACH. As ACHs are paid on a discharge-basis to encourage discharge as early as possible, this may result in hospital-level care being shifted to PAC settings. For example, the types of patients discharged to LTCHs could remain in the ACH and be treated in their step-down unit and later in the general medical or surgical unit in those hospitals, all of which would be included in the initial IPPS MS-DRG payment or qualify as an IPPS outlier.

Medically complex patients that require longer term rehabilitation services and medical treatments, such as patients with spinal cord injuries, traumatic brain injuries, or other neurologic conditions, also cause payment concerns. These patients are commonly admitted to either an LTCH or IRF depending on the availability of different options and the local treatment teams' familiarity with these types of providers. PAC referral patterns have been shown to vary in response to the availability of LTCHs and IRFs in a local market and in response to the whether

a PAC setting is co-located within the acute care hospital². While this arrangement provides greater access to appropriate services, these two types of providers are paid under different prospective systems which recognize different cost components in their payment structures. For example, the LTCH Prospective Payment System (PPS) does not recognize the functional impairments in determining case-mix groups, and the IRF-PPS recognizes only limited severity differences within medical diagnosis, including primary conditions and comorbidities but does not consider other types of medical complicating factors included in the SNF and HH PPS, such as pressure ulcers.

Another patient population of interest is a subset of rehabilitation patients, particularly post-surgical orthopedic and cardiac patients. They may receive post-hospital physical therapy services in one of several sites including IRFs, SNFs, or through community-based services such as home health care or outpatient therapy. However, Medicare payments for the case-mix adjusted patient component may be very different at these sites of care.

1.2 Inconsistencies in Case-Mix Systems and Unintended Incentives

Currently, Medicare uses different PPSs for each of the PAC providers, each with its own case-mix groups, payment units, associated payment rates, and incentive structures. Each of these systems measures case-mix complexity, but each uses a unique set of items to measure the concepts. The current Medicare payment methods for PAC providers are designed largely as independent systems that measure within-setting variation but they do not recognize the potential overlap in case mix or complementary service options available in other settings. More importantly, the variability in case-mix measurement and payment methodologies, including

² Gage, B. et al. ASPE Final Report: Examining Relationships in an Integrated Hospital System. Prepared for Susan Bogasky. <http://aspe.hhs.gov/health/reports/08/examine/report.pdf> . March 2008.

both units and adjustment approaches, makes it difficult to compare patient or facility cost differences in a standard way across settings and to create consistent incentives across payment systems.

The payment systems used in the four post acute care settings are complex and have many nuances and restrictions. In addressing the issue of unintended consequences of independent payment systems, CMS may want to examine a variety of issues and determine whether some payment differences between settings are necessary and warranted. Important considerations in moving towards improved payment systems include consistent measurement of items included in the case mix systems, appropriate representation and detail of patient severity measures within the systems, decisions on whether separate models are necessary for predicting different sub-types of patient specific costs (whether the same or different patient characteristics and weights are used for predicting therapy, routine, and non-therapy ancillaries costs), the appropriate unit of payment (including per diem, per discharge, per episode, or another unit) and whether a common unit of payment should be utilized across settings, in addition to other concerns.

SECTION 2

DATA COLLECTION IN THE DEMONSTRATION

2.1 General Data Collection Approach

Two types of data were collected from the participating providers. All providers, including both acute care hospitals and PAC providers, collected the standardized assessment item set discussed in Section 2.2. Second, PAC providers collected Cost and Resource Use (CRU) data which provided staff time measures for treating each of the enrolled beneficiaries. The primary focus of CRU is staff-time measurement capturing variations in types of staff,

licensure levels, and total time spent with individual patients during three two-week long data collection windows within the nine month CARE collection period in each facility.

Data were obtained in a manner that created a “snap shot” of patient acuity and patient resources within each provider type in proportion to how often that type of patient was seen. This approach provides information on the type of patient treated in each setting at different points in their care trajectory. The sample was designed to be representative of important types of experiences, and thus oversampled certain types of providers and patients in order to capture important clinical and resource expenditure variation. Data were collected from 140 providers in 11 market areas. Market areas and providers were selected to account for the following factors: (1) variation in the types, sub-types and numbers of PAC providers available, including hospital-based versus freestanding providers and chain membership versus independent status; (2) geographic variation, by census region and by urban/suburban/rural populations served; and, (3) beneficiary/patient representativeness capturing typical conditions treated in the Medicare PAC populations and additionally targeting patients most likely to be capable of being treated in multiple settings. Within the market areas, specific numbers of each sub-type of provider were targeted for recruitment based on the characteristics of the market. The number of participating providers, broken down by provider type and market area, is presented in Table 1. The 140 participating providers included 15 acute care hospitals, 15 LTCHs, 26 IRFs, 43 SNFs, and 41 HHAs.

Table 1: Count of participating providers, by provider type, by market area

	Acute	LTCH	IRF Free-stand	IRF Hosp Unit	SNF Free-stand	SNF Hosp Unit	HHA Free-stand	HHA Hosp based	Total
Boston MA	2	4	2	0	7	1	2	2	20
Chicago IL	1	2	2	3	6	0	5	0	19
Columbia MO	1	0	0	0	1	1	5	0	8
Dallas TX	2	3	3	2	4	0	4	1	19
Lakeland/ Tampa FL	1	2	2	0	1	0	3	0	9
Lincoln/Omaha NE	2	0	0	2	3	0	2	2	11
Louisville KY	1	2	3	1	3	1	1	1	13
Rochester NY	1	0	0	1	3	1	5	1	11
San Francisco/ Sacramento CA	2	0	0	2	4	2	2	0	12
Seattle WA/ Portland OR	0	1	0	1	3	0	0	2	7
Sioux Falls SD	0	1	0	0	1	0	0	0	2
Wilmington NC	2	0	0	2	1	1	3	0	9
Total	15	15	12	14	37	6	32	9	140

- Free-Stand: Free standing facility, Hosp Unit: Hospital Unit
- Market areas include providers located within a 2-hour driving radius around the center of each city.

2.2 Development of a Standardized Patient Assessment Tool

No common standardized set of assessment items existed at the beginning of the PAC-PRD initiative. In order to comply with the Congressional directive, CMS developed a uniform assessment instrument that addressed the analytic needs and range of patients seen across the participating provider types: the Continuity Assessment Record and Evaluation (CARE) tool.

The guiding principles for the creation of the CARE tool came from the originating DRA 2005 legislation. These include (1) measuring patient specific factors predictive of resource intensity needs to inform payment policy discussions; (2) measuring patient outcomes and associated risk adjustment; and (3) documenting clinical factors associated with patient discharge placement decisions to provide for seamless and appropriate care transitions.

The CARE tool was designed to collect data at discharge from acute care hospitals and at admission and discharge from each PAC setting. In addition to these assessments, the CARE tool collects data appropriate for use after a significant change in patient status or after a return from an interrupted stay, and a subset of data to be collected in case of patient death. For the collection of time-sensitive data, CARE established standard assessment windows (timeframes) of the first two days following admission and the last two days of a stay prior to discharge.

The collected sample of CARE instruments is shown in Table 2, broken down by assessment form and provider type. The sample contains over 39,000 assessments across all five provider types, including 10,381 assessments in HHAs, 8,996 assessments in SNFs, 10,666 assessments in IRFs, 6,529 assessments in LTCHs. In ACHs, 2,633 discharge assessments were collected.

Table 2: CARE assessment counts by assessment form, by provider type

Overall	HHA	SNF	IRF	LTCH	Acute	Total
Admission	5,116	4,517	5,382	3,141	—	18,156
Discharge	4,489	4,048	5,226	2,751	2,633	19,147
Expired	32	145	10	282	—	469
Interim	744	286	48	355	—	1433
Total	10,381	8,996	10,666	6,529	2,633	39,205

– Dash (—) indicates that these assessment types were not collected at ACH providers.

Minimizing provider data collection burden was a key consideration in the creation of the CARE tool. The estimated provider burden was 30 minutes to one hour per assessment, depending on the complexity of the patient being assessed and assuming that the person completing the tool was trained and familiar with the tool. This time estimate is the same as or

less than the estimated burden of the current CMS assessment tools, such as the MDS and OASIS. If these standardized CARE items replaced analogous existing items in the current hospital intake assessments and the mandated items in the current PAC assessments, minimal or no additional time should be needed to complete a patient intake or discharge assessment relative to currently used assessment items.

CARE provides a standardized set of data elements drawn from existing, extensively tested assessment items that uniformly measure medical, functional, and cognitive status across settings over time. Emphasis was placed on identifying appropriate measurement standards and, to the extent possible, building on current measurement science. The inclusion of self-report and performance-based items was stressed in order to capture the experience of the Medicare beneficiary in addition to the clinician's measurement of the patient's health status.

The CARE item set is comprised of items that are consistent with the domains currently collected in most patient assessment tools or intake assessments and the three CMS mandated tools. Not all CARE items are collected on all Medicare patients. CARE incorporates screener questions to allow less clinically complex patients to be assessed quickly. At the same time, it includes additional items that standardize ways of measuring severity for patients who have medical or functional conditions identified by the screener items.

The final set of items was determined through a process that emphasized extensive stakeholder input. Input was received throughout the development process from clinicians across the range of disciplines treating patients in ACHs and the four PAC settings as well as numerous national professional and provider associations. These experts provided valuable input regarding the types of concepts needed to distinguish severity and the items that best measured those concepts across all settings. The CARE dataset includes:

- *Administrative Items:* Patient demographic information and basic insurance information.
- *Pre-Morbidity Patient Information:* Baseline data on patient's preadmission status and status before the current spell of illness. These are important factors for risk-adjusting outcomes, including the probability of discharge to the community and expected changes in functional abilities.
- *Current Medical Information:* Factors explaining medical or level of care needs. Patients with greater medical complications may need more intensive settings and higher frequency physician and nursing care. The factors in this section are commonly used in current case mix systems, such as diagnosis, comorbidities, procedures, and skin integrity. These items are typically included in a patient's current medical record.
- *Interview Items: Cognitive Status, Mood and Pain:* Patient-centered interview items reflect the voice of the patient. Cognitive, mood and pain items provide important risk adjusters. Observational measures were also included in order to obtain information on these domains where it was not possible to interview the patient.
- *Impairments:* Screening and supplemental items identifying impairments which may impact a patient's functional abilities or otherwise impact a patient's care needs.
- *Functional Status:* The person's ability to perform specified motor tasks, activities of daily living and instrumental activities of daily living. Functional status items serve as both predictors of resource needs and outcomes.
- *Discharge Information:* Patient discharge destination, discharge support needs, and other non-medical, social support factors that might affect placement decisions and which may improve care transitions.

The CARE items' reliability in each setting was evaluated as part of the demonstration.

Practicing clinicians were asked for feedback on the items' use with different types of patients in their respective settings. Additionally, two types of formal reliability tests were conducted. The first used a traditional inter-rater reliability study approach to focus on the reliability of the standardized items when applied to populations in settings other than those for whom the items were originally validated. The second type of test, where assessors in different settings rated uniform 'hypothetical' patients, examined the degree of agreement when items were used by different disciplines in different settings. In addition, the validity of CARE items was assessed

relative to existing items in the legacy tools (MDS, OASIS, and IRF-PAI) and the parsimony of the measurement approach was evaluated.

Additional information on the testing of the CARE tool can be found in the contractor report posted on the CMS website. Key results include:

- The implementation of CARE within the demonstration was successful. All five settings were able to use CARE items to collect information in a consistent and comprehensive manner for their Medicare populations.
- Participant feedback on CARE was generally positive, particularly regarding the functional status items. Therapists consistently commented that the CARE items were easy to use and provided greater specificity for measuring severity and change in function than the items that had been in the MDS 2.0 and OASIS-B in use at the time of the demonstration. Clinicians in all five settings appreciated the use of standard items for measuring pressure ulcers and other medical factors that affect staffing intensity.
- Reliability testing for CARE showed positive results that are consistent with reliability standards used for previous CMS mandated patient assessment instruments.
- Overall, the inter-rater reliability results showed very good agreement on most items. These results suggest that most of the standardized versions of the assessment items have strong reliability within and across settings. Differences across settings were present but each setting still had acceptable levels of reliability within settings, suggesting these items could be used to measure a patient's progress in a standardized way across an episode of care.

SECTION 3

RESOURCE INTENSITY ANALYSIS RESULTS

This chapter discusses models predicting patient specific resource use in the four PAC settings. Differences in resource intensity across settings are important to measure given that different settings may be treating the same type of patient. The use of the CARE tool and the CRU data makes it possible to compare the relative intensity of services provided in each setting. When the data are stratified by case-mix, the resource intensity information will answer the question of whether different settings provide equal treatment intensity for similar patient types.

The resource intensity models presented in this chapter reflect current practices of utilization; they do not necessarily represent what utilization should have occurred.

The models predicting resource use provide the foundation for a revised payment system by assessing the possibility for uniformity in modeling the patient specific component of payment using the same patient acuity measures. A payment system needs to address the variable costs and resource use that are related to the acuity of the patients being treated and also address the fixed expenses that are not related to the specific patient population but instead reflect the costs of running a business and longer term choices such bed size, areas of clinical focus, and management approach.

Methodological Approach: The data used to analyze the Resource Intensity sample includes all patient stays in IRFs, SNFs, and LTCHs with matching CRU, CARE assessment, and claims data. Additionally, it includes all HHA claims episodes with associated CARE assessment data. The sample size for the CRU analysis is shown in Table 3.

Table 3: Number of cleaned CRU patients, patient-days, and days per patient by setting

Setting	Observed Admissions	Observed Patient-days	Mean observed days per patient
All settings	6,705	79,715	11.89
LTCH	728	6,645	9.13
IRF	1,106	8,256	7.46
SNF	800	6,691	8.36
HHA	4,071	58,123	14.28

– CRU = cost and resource utilization. Days per patient for HHAs were based on claims, not CRU data collection as for the other settings.

The basic measure of resource use is the weighted sum of total staff time per individual patient and is derived from the CRU data. Total staff time includes all direct care staff and

support staff directly involved in the care of specific patients. The weights are national average wages for each staff member's occupation and licensure level. From the sample of days with direct CRU collection, models for predicting resource intensity in unobserved days were developed resulting in the creation of stay or home health episode resource intensity measures.

Two Resource Intensity Index (RII) measures were constructed:

- **Routine RII.** Intensity of care provided by routine, non-therapy staff: nursing, nursing aides, respiratory therapy, social work, and case management; and
- **Therapy RII.** Intensity of care provided by therapy staff: physical therapy, occupational therapy, and licensed speech/language pathology therapists, therapy assistants, and therapy aides.

Key Results: Key results of this examination are presented below. A major focus of analysis in this section is an evaluation of the degree to which it makes to move towards increased consistency in payment models between the four PAC settings³. Additional information and discussion of these results can be found in the contractor report posted on the CMS website.

1) The development of Case Mix Systems using uniform definitions and measures of patient acuity between different settings is possible. This can be accomplished with a limited set of common patient acuity items.

Conceptually, three of the four PAC payment systems (SNF, IRF, and HHA) are similar to each other in that they measure medical, functional, and cognitive status of the patient at admission based on an assessment instrument. These systems use this information to construct payment groups. However, operationally, each PPS uses different items to measure patient severity. The absence of uniform case-mix measures has made it difficult for the Medicare

³ There are many ways in which CMS can move towards greater consistency in how patient level variable costs are reflected in payment including, but in no way limited to:

- consistency in whether different aspects of patient costs are modeled together or independently (e.g. the use of separate or combined models for nursing, aides, therapy, drugs, and other non-therapy ancillaries)
- consistency in how different aspects of patient acuity are measured between settings,
- consistency in the impact of a particular acuity score on predicted resource needs,
- consistency in the base rate associated with the model, and,
- consistency in whether day, admission, or another unit of payment will be used.

program to evaluate the extent to which it is paying consistently for cases with similar complexity treated in different settings.

One of the most important contributions of this project is the consistent measures of patient severity across the different PAC settings, allowing for inter-setting comparisons. The value of using the same measures of a concept across settings is that it allows one to determine equivalent costs independent of setting and to measure changes in outcomes for equivalent cases. The movement towards more consistency between payment systems need not require that the different payment models be identical. Having uniform measurements of patient acuity and outcomes in multiple payment systems is a positive step towards understanding differences in a patient's severity, needs, treatment, and outcomes in a consistent manner between settings and helps foster better communication between providers.

This project tested the reliability of uniform items in each setting and showed the items were as reliable as those currently used in the IRF, SNF, and HH PPS. The items tested in each setting were well-received by the clinical staff using the items to measure their patient population complexity. The resource predicting models used in this project all began with measures of patient acuity that were established in a uniform manner across all four PAC settings. Although the *impact* (or weight) of a characteristic may vary from setting to setting, a relatively small set of characteristics were significant predictors. Moving towards the incorporation of CARE tool items in the separate payment systems is a possible means of transitioning to a coordinated PAC payment approach.

2) PAC payment systems can be improved by the inclusion of patient acuity measures that are not used in current payment systems.

The variation in the current case-mix systems is based on historical differences in system development and may not include all types of patient level severity adjustors that are associated

with greater resource needs in that setting. For example, the MS-DRG system used in LTCHs accounts for medical complexity and surgical procedures but fails to account for functional and cognitive complexity. Within LTCHs, certain characteristics available on the CARE tool but not on claims were significantly associated with variation in resource intensity. Significant predictors of routine and therapy resource intensity in LTCHs included information on pressure ulcers, communication disorders, sitting endurance, and functional status. This suggests that existing payment systems can be significantly improved and within-payment group variation in provider margins can be reduced through the addition of selected CARE characteristics.

3) *PAC payment systems can be improved by separately examining and modeling the routine, therapy, and non-therapy ancillary aspects of patient-specific resource use.*

The current PAC payment systems differ in whether they attempt to predict patient specific costs as a whole or whether they break patient specific costs into component parts. This project subdivides patient specific costs into routine care (nursing/aides), therapy, and non-therapy ancillaries. By modeling these components separately, the analysis provides insight into the extent to which each component is important in each of the different settings and whether they are associated with similar or different patient factors. As shown in the contractor report, routine resource intensity and therapy resource intensity were related to different patient acuity predictors, suggesting the importance of examining these resources separately.

Therapy resource intensity models are less driven by case mix differences than are routine intensity models. The explanatory power for the routine models was roughly twice the explanatory power seen in models predicting therapy⁴. The lower explanatory power of the

⁴ MSE-based R² rates for routine models ranged from 0.45 to 0.73 depending on the specific model being examined. Equivalent models predicting therapy intensity ranged from 0.25 to 0.36.

therapy models, compared to the routine models, may indicate that further work is necessary to understand variations in use of therapy resources and to improve these models.

Measures of actual therapy use are used in the current payment systems for HHAs, IRFs, and SNFs in order to establish payment rates, either explicitly (HHAs and SNFs) or through certification requirements due in part to the difficulty in predicting therapy intensity. In general, use measures are considered less desirable than patient acuity measures due to the “game-ability” of the measures; i.e., the possibility that providers could increase reimbursements by inappropriate increases in utilization. The therapy predicting models presented in this report avoid the use of measures of resource use (such as number of visits or minutes of therapy) in the models and rely solely on the functional impairment levels as well as medical and cognitive factors.

Non-therapy ancillaries (NTAs) include such services as prescription medication, laboratory tests, imaging, and x-rays. The cross setting comparison of NTAs is complicated by variation in which specific items are included in each settings’ consolidated billing rules. Utilization patterns are greatly impacted by whether a service is billed separately or combined with a payment bundle, in addition to other factors such as the physical location of the NTA provision.

Costs that do not vary by individual patient are typically considered indirect costs and are referred to as “fixed costs” for the purposes of this report. These costs include capital costs which vary greatly between a hospital, a skilled nursing facility, and a home health agency given their different levels of physical plant needs. Each of the provider types examined in this report have additional non-patient or general administrative costs of managing these facilities independent of the staff needed to provide direct care. These costs are not included in the case-

mix adjustment approach as they are considered to be largely independent of patient specific acuity measures. Instead, these indirect costs reflect the costs of certification and program participation for each respective setting as well as reflecting longer term decisions as areas of clinical sub-focus and management style. These fixed costs will be largely correlated with the type of provider being examined, independent of the individual patients they are treating at a particular time.

4) Multiple approaches to the unit of payment are possible. The choice of payment unit will be largely driven by policy considerations rather than empirical results.

The choice of the unit of payment is a critical decision in the development of a payment system. There are three basic choices: day, stay, and episode. Unit of payment may vary by provider type or may be implemented as a consistent unit of payment across all PAC payment systems. The choice of unit of payment should reflect the extent to which there is either discretion or predictability in patient needs, costs, and appropriate lengths of stay. Additionally, it is important to consider issues of ensuring patient access, especially for patients needing institutional levels of PAC care for an extended period of time, and how the availability of services might be impacted by the unit of payment. The choice of unit of payment must strike a balance between the desire to be cost conscious and the possible inappropriate incentives to shorten the length of stay and discharge the patient prematurely to a less intensive level of care.

In the PAC-PRD contractor report, resource predicting models are assessed at both a stay level and a day level in an effort to provide information that can be used in a flexible manner. The 60-day home health episode was considered to be analogous to an inpatient stay and a home health visit was modeled compared to an inpatient day.

The models presented focus primarily on comparing resource intensity between settings on a stay basis. At the same time, it is important to keep in mind that resource use varies within

the stay. Importantly, a longer length of stay is significantly associated with greater resource use intensity for both routine and therapy services. Daily resource intensity tends to be highest in the first week of an inpatient stay (5 to 30 percent higher than the stay average) and lower late in the stay/episode (beyond the fourth week, about 5 to 15 percent lower than the stay average). A model that uses a consistent stay level unit of payment between settings could explore ways to account for these observed variations of resource use within a stay and to appropriately incentivize appropriate transfer decisions. For example, the possibility of a lump sum payment plus a per diem payment for stays/episodes with lengths beyond some threshold could be explored, particularly for inpatient PAC settings.

5) Evidence supports the potential for development of a common case-mix adjustment system for the three inpatient PAC settings: LTCHs, IRFs, and SNFs. This system would calculate the patient-specific resource expenditures portion of payment in the same manner across settings.

So far in this report, the examination of the potential for improving consistency in PAC payment systems has focused on the successful measurement of patient acuity factors in a uniform manner between settings through the use of the CARE tool, the use of an approach that standardizes the modeling approach to resource use so as to separately predict routine, therapy, and NTA costs, and a focus on consistently predicting resource use on a stay-level for greater comparability between provider types. In addition to these three aspects of consistency, the discussion in this section focuses on also examining the potential for developing resource predicting models that are consistent across multiple PAC settings in the weights and significance levels associated with specific patient acuity measures and in the base resource intensity amounts for the models.

Evidence supports the possibility of developing resource intensity models with consistent measures, weights, base rates, and units of payments for both routine and therapy care for the

three inpatient PAC settings (SNF, IRF, and LTCH) with a model for Home Health that is consistent in acuity measure definitions. The home health resource intensity patterns were found to be significantly different based in part on the fewer hours of services being provided in the home compared to an inpatient setting. The methodology and approach for this analysis is described in this section. The results are further elaborated and discussed in sub-results 5.1, 5.2, and 5.3. A full discussion of the resource utilization analysis can be found in the Contractor Report posted on the CMS website.

In order to examine the potential for site-neutral payment approaches, the results of three sets of models were compared. Note that each of the types of models is intended to predict resource intensity in all four PAC settings. The three models are:

- All-PAC Settings model: consistent in all four setting.
 - This model estimates a single and consistent set of case mix weights and a single base resource intensity amount for all four PAC settings (HHA, IRF, LTCH, and SNF).
 - The All-PAC Settings models are composed of two regression models: (1) a component predicting whether routine or therapy services are used, and (2) a component predicting the amount of services used, conditional on some use.
- HHA-Inpatient PAC Settings model: consistent in three inpatient settings.
 - This model estimates a consistent set of case mix weights and base resource intensity amount for all the three inpatient PAC settings (IRF, LTCH, and SNF). This model allows HHAs to use the same consistent measures of patient acuity as in the other three settings but to have a set of case mix weights and base resource intensity amount that differ from the weights and base amounts that the inpatient PAC settings share.
 - The HHA-Inpatient PAC Settings models are composed of three components: (1) a HHA-only component predicting whether routine or therapy services are used; (2) a HHA-only component predicting the amount of services used, conditional on use; and (3) an inpatient-only component predicting the amount of services used (since all inpatient PAC patient stays received some amount of both routine and therapy services).

Setting-Specific model: not required to be consistent between settings.

- This type of model allows each PAC setting to have its own set of case mix weights and base resource intensity amount. The setting-specific models use consistent measures of patient acuity for each of the different settings, but this model is different from the other two models in that it allows the significance and impact of each measure to differ by setting.
- The Setting-Specific models are composed of five components: (1) a HHA-only component predicting whether routine or therapy services are used; (2) a HHA-only component predicting the amount of routine or therapy services used if use occurred; and (3) separate IRF-, LTCH-, and SNF-specific components predicting the amount of services used (since all inpatient PAC patients received at least some routine and therapy services during their stay)

Within the three models above, three variations were tested in order to determine whether site of care was still an important predictor in understanding resource intensity even after controlling for patient acuity. The three variation examined used 1) only patient acuity measures, 2) only indicators of the site of care, and 3) both patient acuity measures and site of care.

The next three sub-results discuss the results of these models and model variations. It should be noted that model coefficients and the significance associated with specific patient acuity measures produced in this report reflect the association between patient acuity and resource use that arose out of current payment policies and practice patterns, and do not reflect normative standards. In other words, the analysis reflects only what was done, not necessarily what should have been done. In the discussion below, a variety of model summary statistics are presented including mean square error (MSE)-based R-square, McFadden pseudo R-square and predicted-to-actual ratios⁵.

⁵ The R² measures used in the analysis are adaptations of the more commonly reported R² measures used in ordinary least squared regressions. The alternative statistics are required due to the complex nature of the models tested which combine multiple sub-models and examine both dichotomous and continuous dependent variables. In comparing models using the same sample, the MSE R² measures how well a model improves on explanatory power beyond a simple model. The McFadden pseudo R² indicates relative fit. For both statistics, a higher measure suggests better fit. The predicted-to-actual ratios indicate the degree to which a model systematically over- or under-estimates RII for subgroups (in this case, subgroups defined by setting). Ratios above 1.0 indicate

5.1) Using the same acuity measures with the same weights and base rates, consistent models predicting patient specific use of routine services can be created for all the three inpatient PAC settings with minimal levels of over or under estimation. Evidence supports modeling home health routine service use separately from the other PAC settings.

Within the three inpatient PAC settings (SNF, IRF and LTCH), indicators of setting are not statistically significant predictors of routine/nursing stay-total intensity, after controlling for patient acuity. This suggests that a single model can potentially be used in inpatient PAC settings for stay-level payment of routine/nursing services.

All-PAC Setting model. The three variations of the All-PAC Setting model were examined first to determine whether the development of a four setting model was supported by the data. The results of the models showed that patient acuity measures were important predictors of routine resource use but, even after adjustment, the home health setting indicator was a significant predictor of routine resource use. The significance levels associated with the HHA setting indicator suggest that there are small, systematic differences between settings in the All-Setting model which remain after controlling for patient acuity. Specifically, the All-PAC Setting model results suggest that payment adjustors for HHAs may need to be based off of a significantly lower base rate than for other settings.

The continued importance of setting after controlling for patient acuity is also demonstrated through examination of the MSE-based r-square statistics. The All-Setting model using patient acuity measures alone explained 64 percent of the variation in RRII. Including both setting and patient acuity measures increased the explanatory power slightly to 71 percent. The predicted-to-actual ratios for the All-Setting model indicate the degree of miss-fit in predicted resource intensity by comparing the predicted payments based on resource intensity to actual

systematic overestimation of the RII (which would result in overpayment, if the model was used), and ratios below 1.0 indicate systematic underestimation of the RII (resulting in underpayment).

payments. As shown in Table 4, the application of the All-Setting model would result in HHAs being overpaid for routine/nursing care by 252 percent ($3.52 \times 100 - 100$) relative to their true resource intensity. IRFs and LTCHs would be underpaid by about 20 percent, and SNFs would be underpaid by 40 percent.

In summary, results suggest significant levels of miss-fit with the All-PAC Setting model.

Table 4: Ratio of Predicted to Actual Resource Intensity Index by Setting, for Therapy and Routine Stay Level RII Alternative Models

Model	McFadden Pseudo R ²	HHA Ratio	IRF Ratio	LTCH Ratio	SNF Ratio
Routine RII Model Performance:					
All-PAC Settings	0.113	3.52	0.77	0.81	0.59
HHA-Inpatient PAC Settings	0.271	1.00	1.10	0.94	1.01
Setting-Specific	0.293	1.00	1.00	1.00	1.00
Therapy RII Model Performance:					
All-PAC Settings	0.059	1.54	0.82	1.01	0.62
HHA- Inpatient PAC Settings	0.136	1.00	1.01	1.15	0.89
Setting-Specific	0.146	1.00	1.00	1.00	1.00

Table notes:

- All models include all 6,194 patients in the CRU sample, and total sample predicted average RII is set equal to the total sample actual average RII across settings for both resource types.
- See text for a description of the model components present within the All-PAC Settings models, the HHA-Inpatient PAC Settings models, and the Setting-Specific models.
- The McFadden Pseudo R² shown is for the model variation using setting plus patient acuity measures. The Setting-Specific model includes setting indicators by definition. Therefore, this variation is the only one that can be compared across all three models.
- The Setting-Specific models separately predict resources for each setting, so the ratio of the predicted to actual ratios for these models are identically equal to 1.0 since the mean model predicted value equals the mean actual value for each setting individually.

HHA--Inpatient PAC model. Next, the same three variations were examined in the model that separately modeled routine RII in HHAs and in the inpatient PAC settings. In comparing the results of these variations, none of the statistics examined indicated that the inclusion of setting indicators improved the models over patient acuity measures alone. For example, including settings variables improved the explanatory power of the model less than one percent over patient acuity measures alone, indicating that setting did not contribute explanatory power beyond what was already captured in the patient acuity factors. Therefore, separating HH from the three inpatient setting models appears to be a reasonable approach.

When HHAs are separated from the inpatient PAC settings, the predicted-to-actual ratios show that the under- and over-payments for routine services are within 10 percent of the actual value. This suggests that it may be possible to construct a payment approach that models patient intensity needs uniformly across inpatient PAC settings using a common set of case mix weights and a common base resource intensity amount. A separate HH component of the model could be based on consistent measures of patient acuity, but would vary from the inpatient component in both the base rate and the weights assigned to the acuity measures.

Setting-Specific model. The Setting-Specific models use the same patient acuity measures used in all the models presented but here the coefficients associated with the measures are allowed to be set separately for each provider type. The model statistics indicating relative fit were improved only marginally by the use of the Setting-Specific model compared to the HHA-Inpatient model (from 27.1 to 29.3 percent).

Summary of Routine RII model comparisons. The predicted-to actual ratio patterns suggest that a multi-setting model that includes HHAs would be inadvisable. The nature of the service frequency and type of services provided in HHA versus the three inpatient settings are

sufficiently different, so that case mix characteristics have systematically different effects on resource use in HHA than in the other PAC settings. The results of this project suggest that creating a model for paying for variable patient costs for routine services where patient acuity is measured and weighted uniformly in the three inpatient PAC is worthy of further development and exploration.

The Setting-Specific model improves the consistency of payment systems by standardizing the acuity measures but not the weights attached to the measures. This approach preserves the silo based approach in that it allows the clinical measures to have different impacts in each of the payment systems. In the Setting-Specific models, the weight attached to a specific acuity measure (such as a pressure ulcer score) varies by the specific interaction with other factors in that setting rather than being consistent across settings. Using more inclusive models, such as in the HHA-Inpatient setting model, allows for patient acuity measures to be given equivalent weights and thus separates them from some of the setting-specific idiosyncrasies that are reflective of historic practice patterns rather than necessarily being tied to patient needs.

Before a payment model that sets a single set of case mix weights and base routine resource intensity amount across all inpatient PAC settings can be created, more work is needed to refine and simplify the models. After further development work has been done and any remaining issues have been addressed to the furthest extent possible, the policy choice between the desired tradeoff between increased payment system consistency and goodness of model fit will drive whether a setting-specific or the inpatient PAC setting is considered more desirable.

5.2) Using the same acuity measures with the same weights and base rates, consistent payment models predicting patient specific use of therapy services can be created for SNFs and IRFs with minimal levels of over or under estimation. Therapy models that include all three PAC inpatient settings including LTCHs show promise for future development.

As in the analysis of routine service use, the analysis predicting therapy service intensity examined three models: the All-PAC setting model, the HHA-Inpatient PAC setting model, and the Setting-Specific model. Within each of these models, the impact of provider type after controlling for patient acuity was tested.

For models of routine service use, a three setting inpatient PAC approach seems feasible and worthy of further exploration. For models of therapy service use, a two setting model including IRFs and SNF seems most promising as the predicted to actual ratios are within 1 percent and 10 percent, respectively, whereas LTCH predicted values are within 15 percent of the actual value. With refinement, the inpatient model (including all 3 settings) for therapy may also be a possible approach.

All-PAC Setting model. Similar, to what was found in the analysis of routine care, the results of the All-PAC setting models showed that, even after adjustment, the home health setting indicator was a statistically significant predictor of therapy resource intensity. As shown in Table 4 above, the predicted-to-actual ratios for total therapy resource intensity per stay (or 60 day HH episode) show that a payment model based on the All-PAC setting approach would result in HHAs being overpaid for therapy care by 54 percent relative to their true resource intensity. IRFs and SNFs would be underpaid by 18 and 38 percent, respectively. LTCHs' therapy resource intensity is estimated relatively accurately. These predicted-to actual ratio patterns suggest that a multi-setting therapy model that includes HHAs would be inadvisable.

HHA--Inpatient PAC model. When both setting and patient acuity measures were examined in the inpatient-only component of the HHA-Inpatient PAC Setting model, LTCHs were statistically significantly negative versus SNFs, but IRFs were not significantly different than SNFs. This suggests that a therapy payment model combining the inpatient settings but

excluding HHAs may be feasible for IRFs and SNFs, but that the model would need to be modified to better identify LTCHs' lower therapy stay-total.

In the HHA-Inpatient model, the under- and over-payments would be within 15 percent of the true values (see Table 4 above). LTCHs would be overpaid for therapy services, using this model, by 15 percent, and SNFs would be underpaid by 11 percent. However, before an inpatient-only multi-setting model is created, more research is needed on identifying patterns of characteristics that explain the therapy resource intensity differences between inpatient settings.

Setting-Specific Model. The model statistics indicating relative fit were not appreciably improved by the use of the Setting-Specific model compared to the HHA-Inpatient model (from 13.6 to 14.6 percent). Both sets of models showed a better rate of fit than seen in the All-PAC Settings approach (5.9 percent). The final row of Table 4 gives predicted-to-actual ratios for the Therapy Setting-Specific model. As noted above, these ratios are equal to 1.0 since the mean model predicted value equals the mean actual value for each individual setting.

Summary of Therapy RII model comparisons. By definition, the Setting-Specific model shows the best fit. If a small reduction of fit is tolerable in furtherance of the effort to standardize payment methodologies between settings, a standardized IRF-SNF is promising and a standardized inpatient PAC model, including LTCHs, is possible. A therapy model that uses consistent weights across all four PAC settings does not seem advisable.

5.3) Due in part to the nature of home health service provision of care, a payment model combining home health with the other types of PAC providers is not supported by the analysis.

As discussed above, the pattern of resource intensity in the home health setting is significantly different than seen in the three inpatient PAC settings and may require a separate model for explaining resource intensity or a different base rate for payment. This may be due in part to the different nature of providing care and staffing in a home based setting compared to an

inpatient setting. For example, all inpatient PAC stays used at least some routine service use and some therapy service use during their stay. In contrast, home health stays may contain only therapy or only routine, or a combination of services. This nature of service provision will make it more difficult to provide for complete consistency between home health and the inpatient PAC providers. At the same time, a revised home health payment system could be developed using the same underlying definitions of patient acuity measures as used in the other PAC settings even if the exact way that the acuity measures are weighted in the payment system differs. This revised system would use patient acuity measures based on the CARE tool which provides standardized patient assessment information across settings. A revised HHA payment system using CARE measures of patient acuity would provide greater comparability of population acuity, needs, and resource use treated in different settings and at different points in their trajectory of care following a hospital stay in addition to allowing better understanding of cross-setting resource needs and use.

SECTION 4

OUTCOMES ANALYSIS RESULTS

This section presents comparisons of three types of outcomes in the different PAC settings after controlling for patient characteristics: (1) change in self-care functioning from admission to discharge, (2) change in mobility functioning from admission to discharge, and (3) readmission to the hospital within 30 days of discharge from the hospital.

Methodological Approach: Functional ability was assessed on a scale that combines performance for multiple types of activities into an overall measure of ability. The functional scales used range from 0 to 100 with higher numbers indicating more independent functioning. The models used in the functional change analysis predict change in these scales from admission

to discharge within a single PAC setting. All patients included in the examination have both an admission and discharge CARE assessment, so all deaths and emergency readmission patients are excluded. HH patients were assessed at the end of the *last* Medicare-covered 60 day episode before leaving the care of a particular agency. Change in self care functioning and change in mobility functioning are modeled separately. Depending on the condition (i.e., lower amputation versus a neurological condition like multiple sclerosis) the impact of treatment on mobility may be different than the impact on self-care ability. Measuring the two function sub-scales separately allows those differences to be examined.

The dependent variable for the readmission outcome analysis is whether there was an admission to an acute care hospital within 30 days of the previous acute care hospital discharge. All readmissions were counted regardless of reason or whether it was related to the prior care. There was no restriction placed on the location of the patient at this time of readmission. The patient could be in a PAC setting or in the community after PAC discharge. Patients who were readmitted and died within the 30 day window are included but patients who died within 30 days without a readmission were excluded from the analysis.

Key Results: Key results of this examination are presented below. Additional information and discussion of these results can be found in the Contractor Report (posted on the CMS website).

- 1) After controlling for the patient acuity measures, provider type is a statistically significant predictor in the models of change in self care functional ability from admission to discharge.**

Overall, patients had unadjusted self-care change scores that ranged from increases of 9.9 units on the 0-100 scale for the LTCH admissions, to 10.0 for HH admissions, to 12.4 for SNF admission and, finally, to 15.5 for the IRF admissions. LTCH patients had the greatest range in

change scores, ranging from -12.1 to +36.9 points between the fifth and 95th percentiles, compared to IRF patients who had the narrowest range, from +0.0 to +39.2, followed closely by SNF patients, who ranged from -3.6 to +36.7. HHA patients had the highest mean starting self care scores, 59.9 points, compared to 45.4 for SNF patients, 43.6 for IRFs, and 33.9 for LTCH patients.

After controlling for patient acuity at PAC admission, indicators of treatment setting were statistically associated with self care change. This result is discussed in further detail below. At the same time, it is important to note that the multivariate model that controlled only for patient characteristics at PAC setting admission explained 21 percent of the variance in self-care change. Adding setting indicators did not result in a marked increase in the ability of the models to explain change (2 percentage points). This suggests that much of the variation in self-change scores can be explained by patients' medical, functional, and cognitive status at admission.

1.1) IRF stays were associated with a small but statistically significant impact on improving self care functional ability from admission to discharge after controlling for the patient acuity measures. Further exploration is needed to examine whether this impact is related to unobserved factors or therapy intensity in the IRF setting.

After controlling patient acuity differences, IRFs achieved significantly higher changes in self-care scores than the comparison group of SNFs. IRF patients had a mean change measure that was 3.75 units higher than SNF patients ($p < 0.02$). No significant differences in self-care outcomes were present for LTCH patients compared with SNF patients.

The multivariate models controlled for many but not all patient acuity factors that may vary among the populations admitted to each setting. As part of their intake process, IRFs must evaluate and select patients who can tolerate and benefit from three hours a day of therapy at admission. This selection determination may include factors that are associated with greater

improvement but which are not measured in the CARE assessment tool such as patient engagement. Caution is needed in ascribing causation to the preliminary results presented here.

1.2) HHA stays are associated with a statistically significant positive impact on improving self care functional ability from admission to discharge after controlling for the patient acuity measures.

HHA patients had a mean self care change that was 4.02 units higher ($p = 0.001$) than that of SNF patients, after controlling for patient acuity at admission. The clinical importance of these incremental improvements seen in HHAs and IRFs related to this difference has not yet been established. One way of assessing the relative impact of this change is by comparing it to the average unadjusted change in self care in HH: 10.0.

As was noted in the IRF discussion, caution is required in interpreting or assigning causation in response to these results. In interpreting the home health results, it is important to remember that factors such as family involvement that may be associated with both the probability of admission to home health and the amount of self care change were not included in the model. Additionally, the length of time between admission and discharge was not included as a predictive variable and can vary systematically between provider types.

1.3) When examining self care improvement in sub-categories of patients, the significant positive impact of the IRF and HHA settings hold for some but not all diagnosis groups.

Change in self care was examined in selected subpopulations of interest. The musculoskeletal condition group and the nervous system condition group were chosen for examination because of the large number of patients treated in PAC settings for these conditions and because these patients receive a significant amount of physical therapy and/or occupational therapy and may be seen in more than one type of provider.

The musculoskeletal condition group (Major Diagnostic Category (MDC) =8) includes minor and major surgical procedures from the prior acute discharge, spinal diagnoses, and minor and major medical diagnoses. Self-care changes for IRF and LTCH patients were not statistically significantly different from changes for patients treated in SNFs, after controlling for the other variables in the model. The non-significance in IRF for this sub-population stands in contrast to the results for the whole population. Patients with musculoskeletal conditions who received HHA services had higher mean self-care measure changes (4.35 units; $p = 0.02$) than musculoskeletal patients treated in SNFs, and the difference is statistically significant.

The second condition group targeted for separate examination were patients with nervous system conditions (MDC=1). Nervous system patients included patients immediately following a stroke. In the self-care regression model, patients who received HHA and LTCH services had self-care measure changes that were not statistically significantly different than those for patients treated in SNFs, after controlling for patient covariates. In contrast, IRF was significant in the all patient model and the nervous system model. Nervous system patients receiving IRF care achieved a mean change in self-care measure that was 3.93 units higher ($p = 0.02$) than the change for patients treated in SNFs.

2. After controlling for patient acuity measures in CARE, the provider setting is not a significant predictor of change in mobility from admission to discharge.

Prior to adjusting for acuity, patients had mean mobility change scores that varied by PAC setting: +12.1 units gain in mobility for the HH admission, +11.5 for LTCH admissions, +16.6 units for SNF admission, and +16.7 units for the IRF admissions. HHA had the greatest range of scores from the fifth to the 95th percentiles (ranging from -13.0 to +40.3) compared to IRF patients who had the narrowest range over the same percentiles (ranging from +0.5 to +38.6 units).

HHA patients had a mean change that was 2.52 units higher ($p < 0.06$) than that of SNF patients. While slightly higher than the traditional alpha level of .05 for determining statistical significance, the HHA result is noted because it is indicative of a result of possible interest. LTCH and IRF status were not significantly different than SNFs in predicting mobility change. Patient-level covariates explained 21.5 percent of the variance in the mobility change. Adding setting indicators only increased the explanatory power by 0.5%, indicating that adding the provider type variable does not improve the model significantly.

The mobility change score model was examined separately in multivariate models for two high volume conditions of interest: musculoskeletal conditions and nervous system disorders. After controlling for patient acuity, none of the settings were statistically significantly different than SNFs in the mobility changes in either sub-population.

3. After controlling for patient acuity differences, LTCH patients appear to have significantly lower probabilities of being readmitted to an ACH within 30 days of discharge relative to a SNF setting. The interpretation of this result should be made with consideration of the capacity of LTCHs, as hospital-level settings, to deal with higher severity patients.

The overall unadjusted 30 day rate of readmission in the sample was 19.2 percent. IRFs had the lowest, unadjusted, proportion of patients in the sample who were readmitted (17.4 percent) followed by SNFs (19.8 percent), HHAs (20.2 percent) and LTCHs (21.1 percent). The multivariate model for predicting readmission within 30 days explained 4.9 percent of the variation when just patient characteristics at admission to the PAC setting were included. Including setting indicators resulted in only a slight (0.2 percent) increase in explanatory power. After controlling for patient case-mix differences, patients receiving services in HHA and IRFs were similar to SNF patients in their risk for readmission, while LTCH patients had a lower risk (Odds Ratio: 0.56, $p \leq 0.0001$). When readmission was examined in selected populations of

interest, readmission rates were significantly lower for LTCHs for respiratory, circulatory, and nervous system conditions. There were no setting differences among patients with musculoskeletal conditions.

While LTCH patients appear to have a lower risk for readmission within 30 days, it should be noted that, because LTCHs are certified as acute hospitals, the clinical change that would trigger readmission of a patient from an LTCH to an IPPS hospital is different than the clinical change that would trigger a readmission from a SNF to an acute (IPPS) facility. As an acute hospital, LTCHs have more resources that allow them to respond patient changes in medical condition. It is also possible that the differences in the adjusted risk of readmissions by provider type are a reflection of remaining unobserved variation that are simultaneously associated with both readmissions and the likelihood of being discharged to LTCH.

SECTION 5

REVIEW OF RESULTS AND RECOMMENDATIONS

In the initiating legislation, CMS is directed to “establish a demonstration program for purposes of understanding costs and outcomes across different post-acute care sites” and furthermore directs the collection of patient specific acuity, outcomes and costs. In carrying out this mandate, CMS chose to concentrate on standardized measures of the resources needed to care for patients in the different PAC settings and the ability to predict variations in resource use based on measures of patient acuity. These patient specific variable costs are the basis for the development of all of the case-mix adjustment systems and would be the basis of any new system to be developed. Due to the amount of work required for the complete development of a payment system, the focus of this project was on creating a solid basis from which a payment system could potentially be built in the future and to provide information on the extent it seemed

advisable to proceed with development of a system that could cover more than one PAC setting. The project and the RTC contain key information and analyses related to identifying promising areas of future research.

This section re-states the key results presented in this Report to Congress related to the prediction of resource intensity and the prediction of outcomes. Additional models are presented in the Contractor Report posted on the CMS website.

The demonstration was successful in its analytic goals of providing analyses related to the subject matters identified in the initiating legislation and a framework for the consideration of future reform initiatives. Selected additional analyses are already underway or will be completed under related projects using CARE data. Further opportunity for research and development exists beyond these current efforts if desired.

5.1 Key Results in the Resource Intensity Analysis

- The development of Case Mix Systems using uniform definitions and measures of patient acuity between different settings is possible. This can be accomplished with a limited set of common patient acuity items.
- PAC payment systems can be improved by the inclusion of patient acuity measures that are not included in current payment systems.
- PAC payment systems can be improved by separately examining and modeling the routine, therapy, and non-therapy ancillary aspects of patient- specific resource use.
- Multiple approaches to the unit of payment are possible. The choice of payment unit will be largely driven by policy considerations rather than empirical results.
- Evidence supports the potential for development of a common payment system for the three inpatient PAC settings: LTCHs, IRFs, and SNFs. This system would calculate the patient-specific resource expenditures portion of payment using the same acuity measures with the same weights and base rates.
 - Using the same acuity measures with the same weights and base rates, consistent models predicting patient specific use of routine services can be created for all the three inpatient PAC settings with minimal levels of over or under estimation.

Evidence supports modeling home health routine service use separately from the other PAC settings.

- Using the same acuity measures with the same weights and base rates, consistent payment models predicting patient specific use of therapy services can be created for SNFs and IRFs with minimal levels of over or under estimation. Therapy models that include all three PAC inpatient settings including LTCHs show promise for future development.
- Due in part to the nature of home health service provision of care, a payment model combining home health with the other types of PAC providers is not supported by the analysis.
- In the future, CMS may want to consider the refinement of the measures used in the models; the examination of alternative approaches and statistical methods; streamlining of the models by combining measures and other methods; and the examination of predictors of non-therapy ancillary use expenditures and total expenditures.

5.2 Key Results in Outcomes Analysis

- After controlling for the patient acuity measures, provider type is a statistically significant predictor in the models of change in self care functional ability from admission to discharge.
 - IRF stays were associated with a small but statistically significant impact on improving self care functional ability from admission to discharge after controlling for the patient acuity measures. Further exploration is needed to examine whether this impact is related to unobserved factors or therapy intensity in the IRF setting.
 - HHA stays are associated with a statistically significant positive impact on improving self care functional ability from admission to discharge after controlling for the patient acuity measures.
 - When examining self care improvement in sub-categories of patients, the significant positive impact of the IRF and HHA settings held for some but not all diagnosis groups.
- After controlling for patient acuity measures in CARE, the provider setting is not a significant predictor of change in mobility from admission to discharge.
- After controlling for patient acuity differences, LTCH patients appear to have significantly lower probabilities of being readmitted to an ACH within 30 days of discharge relative to a SNF setting. The interpretation of this result should be made with

consideration of the capacity of LTCHs, as hospital-level providers, to deal with higher severity patients.

- In the future, CMS may want to consider refinements of the outcomes models in order to examine additional outcomes measures and develop mechanisms for linking outcomes to payment and other incentive structures.

5.3 CARE tool after the Demonstration

- The implementation of CARE within the demonstration was successful. All five settings were able to use CARE to collect information in a consistent, reliable, and comprehensive manner for their Medicare populations.
- Participant feedback on CARE was generally positive, particularly regarding the functional status items.
- Reliability testing for CARE showed positive results that are consistent with reliability standards used for previous CMS mandated patient assessment instruments. Overall, the inter-rater reliability results showed very good agreement on most items, suggesting these items could be used to measure a patient's progress in a standardized way across an episode of care.
- The development and testing of the CARE tool was undertaken with the assumption that the CARE tool items can and should have a life beyond the demonstration. CMS believes that the standardization of data collection among settings is both possible and desirable for a variety of reasons including more comparable measurement of function and other outcomes, more comparable risk adjustment, and payment modeling. CMS further believes that the collection of patient specific information in hospital settings such as ACHs and LTCHs is advisable.
- The evolution of the CARE tool beyond the PAC demonstration is an iterative and multifaceted process that will continue into the future. CMS has devoted resources towards evolving CARE through a number of current or planned projects including projects focusing on quality reporting, electronic exchange and submission of data, and projects focusing on additional analysis of CARE and CARE item development. Further work will need to occur in the following areas, among others:
 - Continuing to develop quality measures that can be generated from standardized, electronic CARE data for use across the spectrum of patient care settings and provider types,

- Taking the exploratory payment work performed under PAC-PRD and developing formal and implementable payment models through the rule making process and/or future legislative proposals,
 - Expanding the evaluation of CARE based payment models to Medicaid settings and providers,
 - Revise the CARE items in response to the work in the PAC demonstration and other projects, and,
 - Testing whether the CARE tool variables are strong episode-based risk adjusters to inform the larger discussion of payment bundling.
- Given the promise of the CARE tool and the importance of standardizing the collection of information between settings about patient acuity and outcomes, CMS believes that it should pursue its development efforts towards integrating CARE into the reporting requirements for acute care hospitals, SNFs, HHAs, IRFs, and LTCHs.

APPENDIX 1:
SEC. 5008. POST-ACUTE CARE PAYMENT REFORM DEMONSTRATION PROGRAM. DEFICIT REDUCTION ACT OF 2005

Page 52 of 212

(a) *Establishment.*--

(1) **IN GENERAL.**--By not later than January 1, 2008, the Secretary of Health and Human Services (in this section referred to as the "Secretary") shall establish a demonstration program for purposes of understanding costs and outcomes across different post-acute care sites. Under such program, with respect to diagnoses specified by the Secretary, an individual who receives treatment from a provider for such a diagnosis shall receive a single comprehensive assessment on the date of discharge from a subsection (d) hospital (as defined in section 1886(d)(1)(B) of the Social Security Act (42 U.S.C. 395ww(d)(1)(B))) of the needs of the patient and the clinical characteristics of the diagnosis to determine the appropriate placement of such patient in a post-acute care site. The Secretary shall use a standardized patient assessment instrument across all post-acute care sites to measure functional status and other factors during the treatment and at discharge from each provider. Participants in the program shall provide information on the fixed and variable costs for each individual. An additional comprehensive assessment shall be provided at the end of the episode of care.

(2) **NUMBER OF SITES.**--The Secretary shall conduct the demonstration program under this section with sufficient numbers to determine statistically reliable results.

(3) **DURATION.**--The Secretary shall conduct the demonstration program under this section for a 3-year period.

(b) *Waiver Authority.*--The Secretary may waive such requirements of titles XI and XVIII of the Social Security Act (42 U.S.C. 1301 et seq.; 42 U.S.C. 1395 et seq.) as may be necessary for the purpose of carrying out the demonstration program under this section.

(c) *Report.*--Not later than 6 months after the completion of the demonstration program under this section, the Secretary shall submit to Congress a report on such program, that includes the results of the program and recommendations for such legislation and administrative action as the Secretary determines to be appropriate.

(d) *Funding.*--The Secretary shall provide for the transfer from the Federal Hospital Insurance Trust Fund established under section 1817 of the Social Security Act (42 U.S.C. 1395i), \$6,000,000 for the costs of carrying out the demonstration program under this section.