Development and Maintenance of Standardized Cross Setting Patient Assessment Data for Post-Acute Care: Summary Report of Findings from Alpha 2 Pilot Testing

Prepared for

Charlayne Van
Centers for Medicare & Medicaid Services
Center for Clinical Standards and Quality
Quality Measurement & Health Assessment Group
7500 Security Boulevard
Baltimore, MD 21244

Prepared by

Maria Orlando Edelen
Anthony Rodriguez
Sangeeta Ahluwalia
Emily K. Chen
Sarah Dalton
Michael Stephen Dunbar
Jason Michel Etchegaray
Shira H. Fischer
Wenjing Huang
David J. Klein
Jaime Madrigano
Monique Martineau
Steven Martino
Patrick Orr
Jessica Phillips
Francesca Pillemer
Debra Saliba
Shoshana R. Shelton
Cathy D. Sherbourne
Victoria Shier
Regina A. Shih

PR-3588-CMS

Development and Maintenance of Standardized Cross Setting Patient Assessment Data for Post-Acute Care: Summary Report of Findings from Alpha 2 Pilot Testing

by

Maria Orlando Edelen
Anthony Rodriguez
Sangeeta Ahluwalia
Emily K. Chen
Sarah Dalton
Michael Stephen Dunbar
Jason Michel Etchegaray
Shira H. Fischer
Wenjing Huang
David J. Klein
Jaime Madrigano
Monique Martineau
Steven Martino
Patrick Orr
Jessica Phillips
Francesca Pillemer
Debra Saliba
Shoshana R. Shelton
Cathy D. Sherbourne
Victoria Shier
Regina A. Shih

Charlayne Van, COR
CMS Contract No. HHSM-500-2013-13014I

This project was sponsored by the Centers for Medicare & Medicaid Services under contract no. HHSM-500-2013-13014I.

Preface

The Centers for Medicare & Medicaid Services (CMS) contracted with the RAND Corporation to identify and develop standardized items for use in post-acute care patient assessment instruments. RAND was tasked by CMS with developing and testing items within five areas of focus that fall under the clinical categories delineated in the Improving Medicare Post-Acute Care Transformation Act of 2014: (1) cognitive function and mental status; (2) special services, treatments, and interventions; (3) medical conditions and comorbidities; (4) impairments; and (5) other categories.

This report presents results of the Alpha 2 feasibility test of a set of candidate items for assessing some of these focus areas. Conducted between April 2017 and July 2017, the test was the second of two Alpha tests used to assess the feasibility of candidate items. Like the Alpha 1 test, the results of this small-scale feasibility test informed the design of the national Beta test.

This work was sponsored by CMS under contract No. HHSM-500-2013-13014I. The research was conducted in RAND Health, a division of the RAND Corporation. A profile of RAND Health, abstracts of its publications, and ordering information can be found at www.rand.org/health.

Research Team

Project Leadership

Maria Orlando Edelen, Ph.D. RAND Health	Project Director
Barbara Gage, Ph.D. George Washington University	Project Co-Director
Susan Paddock, Ph.D. RAND Health	Lead Statistician
Debra Saliba, M.D. RAND Health	Senior Clinical Adviser

RAND Project Staff

Shoshana Shelton, M.P.H.	Lead Project Manager
Sarah Dalton, M.A.	Project Manager
Liisa Hiatt, M.S.	Project Manager
Sangeeta Ahluwalia, Ph.D.	Task Lead for Training and Co-Lead for Care Preferences
Brian Briscombe, M.A.	Project Support
Emily Butcher	Project Support
Matthew Cefalu, Ph.D.	Co-Lead for Statistics
Emily Chen, Ph.D.	Task Lead for Technical Expert Panels (TEPs)
Michael Dunbar, Ph.D.	Co-Lead for Impairments, Continence, and Pain
Jason Etchegaray, Ph.D.	Co-Lead for the Long-Term Care Hospital (LTCH) Setting Team
Laura Faherty, M.D.	Co-Lead for LTCH Setting Team
Gavin Fahey	Project Support
Shira Fischer, M.D., Ph.D.	Co-Lead for Medication Reconciliation and Co-Lead for the Skilled Nursing Facility (SNF) Setting Team
Wenjing Huang, Ph.D.	Task Lead for Patient-Reported Outcomes Measurement Information System (PROMIS) Items
Noreen Khan	Project Support
David Klein, M.S.	Quantitative Analysis
Kristin Leuschner, Ph.D.	Communications Analyst

Jaime Madrigano, Sc.D.	Task Lead for Subregulatory Public Comment
Rosalie Malsberger, M.S.	Quantitative Analysis
Monique Martineau, M.A.	Communications Analyst
Steven Martino, Ph.D.	Task Lead for Mental Status
Patrick Orr	Project Support
Elizabeth Petrun-Sayers, Ph.D.	Qualitative Data Collection and Analysis
Lauren Pfeifer	Project Support
Jessica Phillips, M.S.	Project Support
Francesca Pillemer, Ph.D.	Task Lead for Care Preferences
Nabeel Qureshi, M.P.H.	Project Support
Rachel Reid, M.D.	Task Lead for the Inpatient Rehabilitation Facility (IRF) Setting Team
Anthony Rodriguez, Ph.D.	Report Writing
Rachel Ross, M.P.H.	Project Support
Cathy Sherbourne, Ph.D.	Task Lead for Cognition
Victoria Shier, Ph.D.	Task Lead for SNF Setting Team and Co-Lead for Impairments and Continence
Regina Shih, Ph.D.	Task Lead for Medication Reconciliation
Elizabeth Sloss, Ph.D.	Quality Measures
Joan Tucker, Ph.D.	Task Lead for Home Health Setting team
Mary Vaiana, Ph.D.	Communications Analyst

Abt Associates Project Staff

Terry Moore, M.P.H., B.S.N.	Principal Associate/Vice President, Health Policy
Donna Hurd, M.S.N.	Senior Associate
Brenda Karkos, M.S.N./M.B.A., RN	Associate-Nurse Researcher
Allison J. Muma, M.H.A.	Principal Associate

Qualidigm Project Staff

Ann Spenard, M.S.N., RN-BC	Vice President/Principal
Lisa Newton, M.S.N., RN-BC	Consultant
Julia Portale, M.B.A., M.P.H.	Consulting Director

Atlas Project Staff

Julia Rollison, Ph.D., M.P.H.	Senior Principal
Jessica Danaux	Consultant
Lauren Whiteman	Consultant

Contents

Tables

Summary

The Centers for Medicare & Medicaid Services (CMS) has contracted with the RAND Corporation to develop standardized assessment-based data elements to meet the requirements of the Improving Medicare Post-Acute Care Transformation (IMPACT) Act of 2014, Section 2(a). The IMPACT Act, Section 2(a), mandates that CMS develop, implement, and maintain standardized patient assessment items for post-acute care (PAC) settings. The four PAC settings are Home Health Agencies (HHAs), Inpatient Rehabilitation Facilities (IRFs), Long-Term Care Hospitals (LTCHs), and Skilled Nursing Facilities (SNFs). The IMPACT Act mandates, at a minimum, standardized items within five clinical categories:

1. functional status, such as mobility and self-care
2. cognitive function and mental status
3. special services, treatments, and interventions (e.g., need for ventilator, dialysis, chemotherapy, and total parenteral nutrition)
4. medical conditions and comorbidities (e.g., diabetes, heart failure, and pressure ulcers)
5. impairments (e.g., incontinence and impaired ability to hear, see, or swallow).

We chose data elements within these categories to be tested because they meet the requirements of the IMPACT Act and because they might support clinical decisionmaking, care coordination, cost reduction, and improved patient/resident and family experiences. To support item-selection activities, we established the following content area–specific work teams:

1. impairments: vision and hearing
2. cognition and mental status: cognitive status
3. cognition and mental status: depressed mood
4. medical conditions: pain
5. other: care preferences
6. other: medication reconciliation
7. impairments: bladder and bowel continence.

Candidate items for standardization under each of the IMPACT Act categories were identified through an environmental scan, which included a literature review, input from the clinical communities serving the PAC populations, and the technical expert panel. These items were then piloted during two feasibility tests. The feasibility tests, referred to as Alpha 1 and Alpha 2, were conducted between August 2016 and October 2016 and April 2017 and July 2017, respectively.

This report presents results of the Alpha 2 feasibility test. The Alpha 2 test included aspects of cognitive status (executive function), care preferences, medication reconciliation, anxiety and behavioral signs and symptoms, and observational assessments of pain, mood, and cognitive status to assess patients/residents who were unable to communicate. The results of these small-

scale feasibility tests will inform a national Beta test designed to determine how well the items perform when implemented in PAC settings.

Methods

Alpha 2 testing was conducted in three markets: one in the Midwest (Chicago, Ill.), one in the South (Houston, Tex.), and one in the West (Denver, Colo.). Providers were eligible to participate in the Alpha 2 test if they were on lists in the Provider of Services file,[1] had case-mix data available, and had sufficient rates of admissions and discharges to reach target assessments during the field period. We also sought to achieve a mix of provider characteristics. The final facility/agency sample included three sites in Denver (one SNF, one IRF, and one HHA), four sites in Houston (one of each setting type), and eight sites in Chicago (two of each setting type).

The Alpha 2 data-collection design used paired assessments of all patients/residents. That is, each assessment was completed by a market-specific research nurse (there were three in total) and a staff assessor from the participating facility/agency. In addition to enabling calculation of interrater reliability (IRR) through the paired assessments, the three market research nurses served as staff trainers and worked closely with participating facilities/agencies throughout the data-collection period to provide support and guidance and to keep track of the progress of facilities/agencies progress in reaching the data-collection goals. The assessor recruitment and training was thus conducted in two phases: The first was recruitment and training of the research nurses, while the second was the training of facility/agency staff in each market.

Data collection took place over ten weeks within the 15 facilities (four HHAs, four SNFs, four IRFs, and three LTCHs) across the three market areas. Data collection spanned from late April 2017 to early July 2017 and was recorded on handheld tablets programmed with the following assessment form "types": (1) admission assessment for communicative patients/residents (target total = 120), (2) planned discharge assessment for communicative patients/residents (target total = 60), and (3) noncommunicative assessment for noncommunicative patients/residents (target total = 120).

For all of the content areas, there were two primary goals for the analysis of the Alpha 2 assessment data: (1) to determine the feasibility of administering the items and (2) to evaluate IRR between pairs of research nurse and facility/agency staff assessors. To evaluate empirical evidence as to feasibility/ease of use, we examined the time spent to complete the items (provided via nurse and staff self-report of the start and end times for each section) and the number of cases in which the research nurse and facility/agency staff left the item missing or indicated "unable to assess/no response." To determine whether items could be completed with acceptable IRR, we calculated the level of agreement between paired assessors' coded item

[1] Eligibility for the Alpha 2 test was determined using the June 2016 Provider of Services file. See CMS, "Provider of Services Current Files," webpage, January 8, 2018.

responses. We used Cohen's kappa for the categorical variables, weighted kappa for the ordinal variables, and Pearson's correlation for continuous variables.[2]

Data from debrief interviews with the research nurses and facility/agency staff were used to supplement the empirical evidence for feasibility/ease of administration. In addition, the data-collection monitoring process yielded useful information about the feasibility of cross-setting field test administration from a logistics perspective.

Results

Overview of Sample and Data Collection

One of the 15 recruited providers (an HHA) did not contribute any assessments, so the final facility/agency sample was 14 providers. Although the total number of paired communicative admission assessments was very close to the target (118 of the targeted 120), there was significant variation in the number of assessments contributed by each provider. However, the total number of paired communicative discharge assessments and noncommunicative assessments did not meet the targets (42 of 60 and 44 of 120, respectively).

There were several factors that might have contributed to the variability in reaching data-collection goals across providers and assessments. Challenges to data collection included the large number of non–English-speaking patients/residents in facilities/agencies, testing interfering with the regular responsibilities of staff, staff vacations and absences, and fewer admissions than anticipated during the data-collection period. There were also fewer opportunities than anticipated to conduct discharge assessments because of long lengths of stay. In addition, research nurses reported that it was difficult to identify and schedule patients/residents for assessment before they were discharged because staff often did not have sufficient advance notice of a discharge. There was also a variety of factors contributing to the difficulty in reaching data-collection goals for the noncommunicative assessments, including providers having few patients/residents who met the eligibility criteria for noncommunicative assessments and patients/residents or families refusing assessment despite initially agreeing to participate.

Cognitive Status

Two data elements were evaluated for assessment of cognitive status: a subset of items from the Developing Outpatient Therapy Payment Alternatives (DOTPA) was completed based on interaction with and observation of the patient/resident, and the Performance Assessment of Self-Care Skills (PASS) was a performance-based assessment. Although assessors reported finding both data elements straightforward to administer, each had long administration times (five to

[2] Jacob Cohen, "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement*, Vol. 20, No. 1, 1960, pp. 37–46.

seven minutes), especially in the LTCH setting, relative to other tested data elements in Alpha 2, except medication reconciliation. Assessors indicated that relevance of the DOTPA may be low for LTCH patients and SNF residents. Many patients in HHAs and IRFs needed assistance completing the PASS tasks. Missing data were prevalent in the PASS assessment.

IRR was highly variable among the DOTPA items, both overall and across settings, with some items showing very low agreement (as low as 0.34) and others showing excellent agreement (as high as 0.81). In contrast, IRR was consistently high overall for PASS (ranging from 0.78 to 0.92). In comparison to the Brief Interview for Mental Status (BIMS), which is currently used to assess cognitive status in the Minimum Data Set (MDS), both the DOTPA and PASS appear to confer additional information. It may be possible to address issues with the length of time to administer and the need for patient/resident assistance through assessor training, but consideration should be given to whether other limitations of these data elements warrant caution for cross-setting standardization.

Behavioral Signs and Symptoms

The Behavioral Signs and Symptoms data elements assessed the frequency and presence of behavioral symptoms and rejection of care. Each took less than three minutes to complete. Few patients/residents exhibited symptoms that warranted continuing with follow-up items assessing impact of behaviors on the patient/resident and on others, leaving very small sample sizes with little response variability for these items and thus making IRR calculation difficult. Where calculable, IRR was moderate to good (0.60–0.77). The data-collection protocol used for the Alpha 2 feasibility test required assessment within four days of admission, but these data elements use a look-back period of seven days, which presented some difficulty by requiring knowledge about behaviors occurring prior to admission. This look-back period may warrant reconsideration in future testing.

Anxiety

The 11 interview-based items in the Patient-Reported Outcomes Measurement Information System (PROMIS) Anxiety bank represent the first attempt at cross-setting assessment of anxiety in PAC settings. The completion rate for these items was high, with little missing data. Assessors reported finding the instructions to be clear, and they completed the set of items within four minutes, on average. IRR was nearly perfect, ranging from 0.80 to 1.00, and little variation was observed between settings. The look-back period of seven days may need to be reevaluated, because it required patients/residents in this feasibility test to remember their feelings before admission to their current facility/agency. Further, feedback indicated that the patients/residents found the set of items somewhat burdensome and repetitive.

Care Preferences

Assessors gathered information on Physician Orders and Goals of Care documented in the medical records of patients/residents. The rate of completion was high, with low levels of missing data in terms of questions answered. Assessments in this category took four to seven minutes to complete, on average. However, an estimated 75 percent of patients/residents had no documentation of Physician Orders, such as do not resuscitate (DNR) or do not intubate (DNI). Where calculable, IRR for this data element ranged widely, with overall kappas ranging from 0.22 to 0.66. Assessors found it difficult to find information in medical records to complete the Goals of Care data element. IRR for Goals of Care was poor, ranging from –0.35 to 0.49.

Medication Reconciliation

The medication reconciliation data elements were developed to assess whether and how medication reconciliation was conducted. Although this data element appeared feasible to complete in all four settings, there were challenges. Assessments in this category took an average of 12 minutes. Evidence for IRR was mixed, but was unacceptably low for the indications and discrepancies items. IRR for whether a patient was taking any medications within a class ranged from 0.33 to 0.88. IRR for the indications ranged from –0.50 to 0.73 overall, and the IRR for the discrepancies data elements ranged from –0.05 to 0.38 overall. These items require further testing; they were only completed on the subset of patients/residents receiving each of the ten medication classes, and, therefore, the sample size was small.

Noncommunicative Assessments

Three data elements developed specifically for use with noncommunicative patients/residents were included in Alpha 2. Each assessment was conducted via staff observation. The assessments are of pain (Observational Assessment of Pain), mood (Staff Assessment of Patient/Resident Mood), and cognitive status (Staff Assessment of Mental Status).

The Observational Assessment of Pain took approximately three minutes to complete. Assessors' comments generally reflected that the pain items were straightforward but somewhat challenging to administer because of the time required for observation and the need to consult multiple data sources. IRR on observational pain items was substantial to almost perfect.

The Staff Assessment of Patient/Resident Mood does not seem to be overly burdensome, taking an average of 5.5 minutes to complete. However, assessors noted that some items were difficult for staff to assess in LTCHs because they were not applicable (e.g., poor appetite or overeating) or because they were unable to assess inner thoughts without any communication. The small number of HHAs included in this test makes it difficult to draw a conclusion about feasibility in that setting. IRR was high for all items and varied little across settings.

The Staff Assessment of Mental Status appears to be feasible to administer in all settings, taking an average of 3.9 minutes to complete. There were several patients/residents with missing

data or responses of "unknown or unable to assess," indicating that the assessors had some difficulty determining answers to some questions for certain patients/residents. IRR tended to be substantial, except in IRF settings, where reliability was lower for some items.

Conclusion and Recommendations

Findings demonstrate that data elements tested in Alpha 2 were generally feasible to administer, although some data elements appear to be more feasible in certain PAC settings than others. For example, the cognitive status data elements were identified as potentially problematic to administer in LTCHs. Testing on a larger sample will be needed to understand more about the challenges in LTCHs.

Feedback from assessors indicated that instructions for most items were clear and facilitated successful completion. Low levels of missing data for many of the items support this conclusion. While data elements proved feasible to administer, the level of burden undertaken to complete the assessment varied across data elements and settings. Time to complete data elements varied widely, with Behavioral Signs and Symptoms taking less than three minutes and Medication Reconciliation taking 12 minutes, on average, to complete. Enhancements to assessor training may be useful to reduce the amount of time it takes to complete some data elements.

Across several data elements, look-back periods that extend prior to the date of admission were sometimes difficult for the patient/resident to recall and may warrant reconsideration in future testing. The optimal look-back window for these data elements is worthy of further attention to identify the best specification for cross-setting standardization. This is particularly challenging in light of the varied reporting requirements that exist across the four PAC settings.

Overall, data elements tested in Alpha 2 exhibited good interrater agreement. However, reliability was difficult to assess for some data elements, including several of the behavioral items, because of the very small number of behaviors that were exhibited, and many Care Preferences items, because of the infrequency with which Physician Orders were found in medical records. A larger sample size in future testing will enable calculation of IRR for the items and data elements that lacked data in this feasibility test. The reliability for Medication Reconciliation was mixed. For many of these items, IRR was moderate to high. For others, IRR was unacceptably low.

Results from the Alpha 2 test inform several recommendations that should be considered to maximize provider recruitment and assessment completion in any future testing. In the recruitment phase, recruiters should clearly convey field test requirements to all potential provider participants so that providers that cannot support data collection (e.g., providers with few English-speaking patients) are not enrolled. Recruitment outreach to potential provider participants would also benefit from enhanced descriptions of the project and objectives for testing. With respect to training, research nurses should be made aware of the importance of building positive engagement with their facility/agency staff partners. The difficulty in flagging

and completing discharge assessments should also be addressed through improved research nurse training on processes to identify and communicate information on planned discharges to field staff. Field staff training should be enhanced with additional examples of ways to introduce and explain the assessments to different types of patients to encourage participation. Finally, training should be an ongoing activity; research nurses and facility/agency staff should be encouraged to conduct initial rounds of the assessments together, and project staff should be prepared to conduct periodic educational updates to make sure assessments are completed consistently throughout the data-collection period across all participating providers.

Acknowledgments

We wish to acknowledge the insightful guidance and input received from staff at the Centers for Medicare & Medicaid Services, including Stella Mandl, Mary Pratt, Tara McMullen, Teresa Mota, and Charlayne Van.

We are grateful for the contributions of the RAND, Abt, and Qualidigm team members listed on pp. iv and v. We thank the research nurses, Cindy Lippman, Robin Richardson, P. J. Lucas, Kathleen Ruroede, Trinette Lindsey, and Antrelene Ward, and the facilities in Denver, Houston, and Chicago that participated as field test providers. We also extend our appreciation to the wide variety of stakeholders, especially to the advisers and technical expert panel members who provided input and perspectives that guided selection of the data elements to include in the field test.

Finally, we thank Justin Timbie of RAND and Barbara Resnick of the University of Maryland for their thoughtful reviews of the report.

Abbreviations

ADL	activities of daily living
BIMS	Brief Interview for Mental Status
CARE	Continuity Assessment Record and Evaluation
CMS	Centers for Medicare & Medicaid Services
DNI	do not intubate
DNR	do not resuscitate
DOTPA	Developing Outpatient Therapy Payment Alternatives
EHR	electronic health record
HHA	Home Health Agency
IMPACT Act	Improving Medicare Post-Acute Care Transformation Act of 2014
IRF	Inpatient Rehabilitation Facility
IRF-PAI	Inpatient Rehabilitation Facility–Patient Assessment Instrument
IRR	interrater reliability
LCDS	Long-Term Care Hospital Continuity Assessment Record and Evaluation Data Set
LTCH	Long-Term Care Hospital
MDS	Minimum Data Set
MR	medication reconciliation
N/A	not applicable
NIH	National Institutes of Health
OASIS	Outcome and Assessment Information Set
PAC	post-acute care
PASS	Performance Assessment of Self-Care Skills
PHQ	Patient Health Questionnaire
PROMIS	Patient-Reported Outcomes Measurement Information System

ROC	resumption of care
SD	standard deviation
SNF	Skilled Nursing Facility
SOC	start of care
TEP	technical expert panel

Chapter One. Introduction

Project Overview

The Centers for Medicare & Medicaid Services (CMS) has contracted with the RAND Corporation to develop standardized assessment-based data elements to meet the requirements of the Improving Medicare Post-Acute Care Transformation (IMPACT) Act of 2014, Section 2(a). The contract name is "Development and Maintenance of Post-Acute Care Cross-Setting Standardized Assessment Data." The contract number is HHSM-500-2013-13014I.

The IMPACT Act, Section 2(a), mandates that CMS develop, implement, and maintain standardized patient assessment items for post-acute care (PAC) settings. The four PAC settings are: Home Health Agencies (HHAs), Inpatient Rehabilitation Facilities (IRFs), Long-Term Care Hospitals (LTCHs), and Skilled Nursing Facilities (SNFs). Existing PAC assessment instruments by setting are Outcome and Assessment Information Set (OASIS) for HHAs, Inpatient Rehabilitation Facility–Patient Assessment Instrument (IRF-PAI) for IRFs, LTCH Continuity Assessment Record and Evaluation (CARE) Data Set,[1] and Minimum Data Set (MDS) for SNFs. Standardized items are to be nested within the four existing PAC assessment instruments; however, each instrument will continue to have unique items selected for their special relevance to their respective PAC settings. The IMPACT Act mandates, at a minimum, standardized items within the following clinical categories:

1. functional status, such as mobility and self-care
2. cognitive function and mental status
3. special services, treatments, and interventions (e.g., need for ventilator, dialysis, chemotherapy, and total parenteral nutrition)
4. medical conditions and comorbidities (e.g., diabetes, heart failure, and pressure ulcers)
5. impairments (e.g., incontinence and impaired ability to hear, see, or swallow).

Again, in consultation with CMS, we selected data items within these categories because they meet the requirements of the IMPACT Act and because they might support clinical decisionmaking, care coordination, cost reduction, and improved patient/resident and family experiences. To support item-selection activities, we established the following content area–specific work teams:

1. impairments: vision and hearing
2. cognition and mental status: cognitive status
3. cognition and mental status: depressed mood
4. medical conditions: pain
5. other: care preferences

[1] The LTCH CARE Data Set is abbreviated as LCDS.

6. other: medication reconciliation
7. impairments: bladder and bowel continence.

In addition, we established a cross-category work team to consider cross-setting standardization efforts from the perspectives of workflow, interoperability, and care transitions.

Each work team was led by RAND researchers and included advisers, clinicians, and academic researchers with expertise in PAC settings. RAND staff led the research activities but actively collaborated with clinical and academic advisers on an ongoing basis. Work teams were overseen by project leadership: Project Director, Maria Edelen, Ph.D. (RAND), and Project Co-Director, Barbara Gage, Ph.D. (George Washington University), with clinical content support from Debra Saliba, M.D., M.P.H. (RAND). The lead statistician in this effort was Susan Paddock, Ph.D. (RAND). Sangeeta Ahluwalia, Ph.D. (RAND), led assessor training, and Emily Chen, Ph.D. (RAND), coordinated such key stakeholder activities as technical expert panels (TEPs) and public comments.

Candidate items for standardization under each of the IMPACT Act categories were identified through an environmental scan, which included a literature review, consultation with experts in the field, input from the clinical communities serving the PAC populations (e.g., focus groups, public comments), public comment periods, discussions with stakeholders, discussions with partners within CMS and the U.S. Department of Health and Human Services, and feedback from our TEP. These items were then piloted during two feasibility tests. The feasibility tests, referred to as Alpha 1 and Alpha 2, were conducted between August 2016 and October 2016 and April 2017 and July 2017, respectively. The data elements tested in Alpha 2 included items from five of the six RAND categories. (Alpha 2 did not include any data elements in the impairments category because they were tested in Alpha 1 and did not require retesting for feasibility.)

The results of the Alpha 1 test, together with the Alpha 2 test discussed in this report, are informing a national Beta test to evaluate candidate data element performance when used in any of the four PAC settings. Unlike alpha feasibility testing, beta testing involves a much larger sample size, because the intent is to provide sufficient power for hypothesis testing and for more-precise estimates of item performance.

The purpose of this report is to describe the methods and results from the Alpha 2 feasibility test.

Organization of This Report

The remainder of this report is organized as follows:

- Chapters Two and Three present the sampling design, recruitment, data-collection, and analytic methods used to conduct the Alpha 2 feasibility test.
- Chapters Four through Nine report the results, both overall and for data elements in the Alpha 2 test.
- Chapter Ten provides some concluding remarks.

- Appendixes A, B, and C contain background information on data elements and protocols used in the field test.

Chapter Two. Methods

In this chapter, we describe the methods used in conducting the Alpha 2 feasibility test. The chapter is divided into five sections:

- Selection and Development of Data Elements for Testing
- Market and Facility/Agency Selection and Recruitment
- Assessor Recruitment and Training
- Data Collection
- Analytic Approach.

Selection and Development of Data Elements for Testing

Data elements for consideration in Alpha 2 testing were identified after a rigorous review and development process. In particular, the Alpha 2 test protocol included three data elements to assess noncommunicative patients specifically and new or modified data elements within the categories of Cognition and Mental Status, Care Preferences, and Medication Reconciliation (MR). A final list of candidate data elements for Alpha 2 testing was determined in consultation with CMS and with guidance from a January 2017 meeting of the TEP.

A summary table of the data elements tested in Alpha 2 can be found in Table A.1 in Appendix A. The table covers whether each group of items is currently in use and in which PAC assessments, what evidence exists for the feasibility and reliability or validity of the items (if any), and the mode of data collection.

Market and Facility/Agency Selection and Recruitment

Alpha 2 testing was conducted in three markets: one in the Midwest (Chicago, Ill.), one in the South (Houston, Tex.), and one in the West (Denver, Colo.). These three metropolitan areas were selected to ensure geographic variation in Alpha 2 testing and to reflect regional differences in practice patterns. Markets encompass the area within a 1.5-hour drive-time radius of the target city's downtown and within the hospital referral region for each of these metropolitan areas.

We sought to recruit at least one facility/agency of each PAC type in each of the three markets. To offset potential challenges to reaching our overall assessment completion goals across the three markets, we sought to recruit two PAC facilities per setting type (rather than one) in the Chicago market. Chicago was selected to contribute additional facilities because it is the largest market in the Alpha 2 test and could most easily accommodate the increase. The final facility/agency sample included three providers in Denver (one SNF, one IRF, and one HHA), four providers in Houston (one of each setting type), and eight providers in Chicago (two of each setting type).

Information from Medicare administrative files was used to choose providers for participation in the Alpha 2 test. To be eligible, providers needed to be on lists in the Provider of Services file.[2] The file contains an individual record for each Medicare-approved provider and is updated quarterly. The data include characteristics of hospitals and other types of health care facilities, including name and address. The data are collected through the CMS Regional Offices. In addition, case-mix data needed to be available for all providers. We obtained these data for LTCH providers from the fiscal year (FY) 2016 LTCH Final Rule Impact File and for IRF providers from the FY 2016 IRF-PAI Final Rule Impact File.[3] The Impact Files contain the universe of LTCHs and IRFs used to estimate policy updates in the final LTCH and IRF prospective payment system regulations and include variables on the case-mix index, the annual number of cases per facility/agency, ownership status, location, and other facility/agency–level variables. Case-mix data for SNF and HHA providers came from CMS's utilization and payment public use files, which were released in 2016, with data covering calendar year 2013. Five-star quality ratings for HHAs were obtained from the CMS utilization files, and the ratings for SNFs and SNF ownership status were obtained from the Nursing Home Compare data archive.[4] Rehabilitation Impairment Categories (RICs), which characterize case mix for IRFs, were obtained from the 2013 IRF-PAI. Summaries of the distribution of diagnosis-related groups (DRGs) at LTCHs were obtained directly from CMS.

Providers were selected to represent a mix of the following provider characteristics: number of admissions per year, hospital affiliation, rural/urban location, profit status, and case-mix distribution. In addition, maps of the geographic locations of each potential provider by provider type and market were developed by the RAND team to select eligible providers that were near one another to increase the data-collection efficiency of the research nurses working in each market.

To ensure a sufficient number of assessments for meaningful data analysis, targets were set for numbers of assessments at the four types of PAC facilities/agencies. Recruitment efforts for inclusion in the study, therefore, were focused on PAC facilities/agencies that were large enough to support data collection (i.e., had sufficient rates of admissions and discharges to reach the targeted number of assessments during the field period)—specifically, IRFs and LTCHs with at least 100 discharges annually, SNFs with at least 100 annual total stays, and HHAs with at least 100 episodes annually.

Facility/agency recruitment lists, generated from the processes described earlier, were provided to study recruiters. In cases where recruitment efforts yielded insufficient participation,

---

[2] We used the June 2016 Provider of Services file to determine eligibility. See CMS, "Provider of Services Current Files," webpage, January 8, 2018.

[3] CMS, "LTCHPPS Historical Impact Files," webpage, February 29, 2012; CMS, "Inpatient Rehabilitation Facility PPS: Data Files," webpage, August 8, 2017.

[4] Data.Medicare.gov, "Home Health Compare Data Archive," webpage, undated(a); Data.Medicare.gov, "Nursing Home Compare Data Archive," webpage, undated(b).

additional facilities/agencies were added to the list(s). These additional facilities/agencies were either drawn directly from the RAND team's master list of eligible PAC providers in the three markets or could be generated via referral from PAC industry contacts and then verified against the master list. For example, when recruiters were unable to obtain sufficient HHA participation in the Denver market, a home health corporate chain could offer a chain-affiliated HHA that met assessment volume and other sample criteria.

Assessor Recruitment and Training

The Alpha 2 data-collection design used paired assessments of all patients/residents. That is, each assessment was completed by a market-specific research nurse (there were three total) and a staff assessor from the participating facility/agency. In addition to enabling the calculation of interrater reliability (IRR) through the paired assessments, the research nurses served as staff trainers and worked closely with participating facilities/agencies in their markets throughout the data-collection period to provide support and guidance and to keep track of facilities'/agencies' progress in reaching the data-collection goals. The assessor recruitment and training was conducted in two phases: recruitment and training of the research nurses, followed by facility/agency staff training in each market.

Research Nurse Recruitment and Training

The RAND team, in partnership with Qualidigm, recruited research nurses who had experience in at least one of the four PAC settings and who lived in one of the three market areas but were not affiliated with the Alpha 2 test facilities. Qualidigm interviewed and hired registered nurses (RNs) who had been prescreened by its partner agency, Ready Nurse. The primary criterion for selecting research nurse participants was that the nurse must have had direct experience administering and completing the assessment instrument for that setting (i.e., the MDS in SNFs, OASIS in HHAs, IRF-PAI in IRFs, and LCDS in LTCHs). Six research nurses were hired and participated in data collection: one research nurse in Denver, two in Houston, and three in the Chicago area. They were distributed across the markets based on the sample size of facilities (i.e., the largest market had the most research nurses).

The selected research nurses participated in a "train-the-trainer" program, which was divided into three components: (1) a pretraining webinar completed one week prior to the training to provide relevant study context and prepare nurses for the in-person training activities; (2) a weeklong in-person training held at RAND's headquarters in Santa Monica, Calif., during the week of March 20, 2017, on the administration of the data elements and hands-on practice with the data-collection instrument and electronic data-collection procedures; and (3) a post-training webinar completed one to two weeks after the in-person training to prepare nurses for the field training of facility/agency staff. Upon completion of the training, the research nurses returned to

their market areas and assisted the RAND team in trainings for all facility/agency staff who participated in Alpha 2 data collection.

The RAND team also created a user manual to accompany the Alpha 2 data-collection instrument, which specified how the items were to be collected.

Facility/Agency Staff Training

The RAND team conducted one field training in each of the three markets in a centrally accessible Alpha 2 test facility/agency or other location with conference capabilities. Field trainings were held during the weeks of April 17, 2017, and April 24, 2017. A total of 38 staff attended one of these trainings (21 in Chicago, ten in Denver, and seven in Houston). Each field training was 1.5 days in duration, with the first day focused primarily on training facility/agency staff to administer each data element. The following half day focused on scripted role-play exercises designed to give participants hands-on practice with the data-collection instrument and electronic data-collection process.

Each field training was led by a "core" of three trained research nurses (one from each market) and was supported by the remaining local research nurses from that market. This research nurse core traveled to each market to deliver the field trainings. In addition, each field training was attended and supported by key project personnel, including members of the RAND, Abt, and Qualidigm teams. Prior to the first field training, key project personnel and research nurse trainers conducted a day-long dry run of the field training to smooth out logistics, presentation delivery, and other issues.

Data Collection

Alpha 2 data collection took place over ten weeks within the 15 facilities (four HHAs, four SNFs, four IRFs, and three LTCHs) across the three market areas specified earlier. Data collection spanned from late April 2017 to early July 2017. Each completed Alpha 2 assessment comprised two assessment forms to evaluate IRR: one from a research nurse and one from facility/agency staff.

The Alpha 2 data collection was designed to provide performance results for several standard assessment items meant to be suitable for all patients/residents who are able to communicate (henceforth referred to as *communicative assessment*), and to evaluate the performance of three observational assessment data elements meant to be administered to patients/residents who are unable or unwilling to communicate meaningfully (henceforth referred to as *noncommunicative assessment*). To accommodate this need, the overall target assessment completion goals across all three markets were set at 120 communicative admission assessments, 60 communicative discharge assessments, and 120 noncommunicative assessments. To increase the likelihood of completing noncommunicative assessments, these assessments were not tied to an admission or

discharge date. We defined one completed assessment as an assessment done by both a research nurse and a facility/agency staff assessor.

Each Alpha 2 PAC facility/agency was asked to complete ten communicative admission assessments, five communicative discharge assessments (among patients/residents for whom a communicative admission assessment had already been conducted), and ten noncommunicative assessments over the ten-week data-collection period. (See Appendix B for the Alpha 2 assessment protocol.)

To begin scaling up the data-collection effort for the Beta test, we moved from paper data collection (used in Alpha 1) to electronic data collection in Alpha 2. Electronic data collection provided the opportunities to (1) collect data in a timely and efficient manner across more facilities and data collectors and (2) directly transfer data from the field to the RAND server, ultimately improving data security and efficiency. In Alpha 2, all data collectors were given a handheld tablet programmed with the appropriate candidate data elements to be used in the field. Training modules included separate sections on tablet use, electronic data collection, and electronic data security and monitoring. In addition, practice sessions conducted during the trainings included the use of tablets to ensure that data collectors had adequate time to prepare for electronic data collection in the field.

Tablets were programmed by RAND team members experienced in the development of applications for pilot-test data collection in health care settings. Each tablet was programmed with the following assessment form "types": (1) admission assessment for communicative patients/residents, (2) planned discharge assessment for communicative patients/residents, and (3) assessment for noncommunicative patients/residents.

Data Security

Assessment data were encrypted and uploaded to a secure RAND server each night. Once uploaded and verified, completed assessment forms were wiped from data collectors' tablets to ensure security during the data-collection process. Data collectors tracked assessments through a unique assessment ID tied to patient/resident name, date of birth, and gender. An encrypted crosswalk linking the patient-identifiable information with the assessment ID remained on each data collector's tablet through the ten-week data-collection period. At the end of the test period, the crosswalk was uploaded to the secure RAND server and wiped from the data collectors' tablets.

Debrief Interviews

As a crucial part of evaluating the candidate items during the feasibility test, structured debrief group interviews were conducted via telephone by RAND staff to gather information from the research nurse and facility/agency staff assessors. These interviews provided information on the assessors' general experience of administering the assessment, including the clarity of the assessment materials (e.g., instruction manuals, assessment form), ease of use, the

ability of patients/residents in each setting to understand the questions, nurse and facility/agency staff level of comfort in asking the questions, and potential challenges to assessment administration.

The RAND team conducted three postassessment structured debrief interviews of research nurses and four structured debrief interviews of facility/agency staff involved in the Alpha 2 feasibility test. Research nurse interviews were conducted at the end of week two of data collection, at the end of week five, and after data collection was completed. To maximize participation, facility/agency staff were offered two opportunities to participate in interviews, at the end of weeks three and seven. Interviews were conducted at various points throughout the data-collection period to identify and address early challenges and to learn about challenges that persisted throughout data collection or that came up as data collection progressed. (See Appendix C for the assessor debrief interview protocol.)

Monitoring

The RAND team monitored the data-collection process throughout Alpha 2 feasibility testing. Weekly conference calls between RAND researchers, Qualidigm, and research nurses were held to discuss progress, address problems that arose during data collection, and make clarifications to the data-collection process as needed. Research nurses kept detailed tracking sheets of completed assessments, which were shared with RAND team members and Qualidigm during the weekly monitoring calls and were also reported to CMS at weekly meetings. The RAND team also generated reports of uploaded assessments, which were cross-checked with the research nurse data. Semistructured interviews with the research nurses were also a crucial part of the data-monitoring effort; they helped to reveal challenges with the data collection in a timely and efficient manner.

Help Desk

The RAND team established a help desk website to aid the research nurses and the facility/agency data collectors. The website provided links to data-collection resource materials, including copies of forms, the data-collection manual, data-management tools, and contact information for each of the research nurses. Qualidigm staffed the help desk, in consultation with four PAC-setting expert teams at RAND, and answered questions within two business days of receipt. RAND researchers and Qualidigm tracked the questions and responses, which were entered into a database by a help desk staff member from Qualidigm. The CMS IMPACT Act help desk was also continuously available; questions pertaining to the Alpha 2 test were responded to by CMS in consultation with the RAND team.

Analytic Approach

Data from the communicative admission assessments and the noncommunicative assessments were used to evaluate the feasibility of the data elements for cross-setting standardization, which we describe in more detail in the next section. The purpose of the communicative discharge assessment was to become familiar with the logistics associated with completing both admission and discharge assessments on the same patient/resident in this field-test context, to understand the challenges, and to refine our estimates of the expected number of discharges per admission in a given study period to inform planning for the Beta testing period.

Feasibility of Administration and IRR

For all of the content areas, there were two primary goals for analysis of the Alpha 2 assessment data: (1) to determine the feasibility of administering the items and (2) to evaluate IRR between pairs of research nurse and facility/agency staff assessors. For each set of data elements discussed in the following chapters, feasibility evidence is presented first (as goal 1), and IRR is presented second (as goal 2). For both goals, we used a similar approach for all items. In some content areas, because of the complexity of the items tested, additional goals for analysis of the assessment data are included, such as determining the fidelity with which skip patterns were followed and observing whether the information sources used to complete the items differed among assessors. When other goals were present for a particular set of data elements, they are identified as goal 3 and higher.

To evaluate empirical evidence as to feasibility/ease of administration, we examined (1) the time spent to complete the items (provided via nurse and staff self-report of the start and end times for each section) and (2) the number of cases in which the research nurse and facility/agency staff left the item missing or indicated "unable to assess/no response."

To determine whether items could be completed with acceptable IRR, we calculated the level of agreement between paired assessors' coded item responses. We used Cohen's kappa for the categorical variables, weighted kappa for the ordinal variables, and Pearson's correlation for continuous variables.[5] Magnitude of agreement was determined by conventional criteria.[6] Cohen's kappa is computed from a 2 × 2 table with two agreement and two disagreement cells. When there is no disagreement, resulting in empty cells, kappa is undefined and cannot be computed. In such cases, IRR is not evaluated and is noted as such in this report. Similarly, for instances of sparse data, IRR is not meaningfully interpretable and is thus not reported.

To explore setting-by-setting variation in the feasibility of administration and reliability of the assessments, we calculated all statistics overall and by PAC setting. However, the Alpha 2

---

[5] Jacob Cohen, "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement*, Vol. 20, No. 1, 1960, pp. 37–46.

[6] J. Richard Landis and Gary G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, Vol. 33, No. 1, 1977, pp. 159–174.

feasibility test was not designed to produce robust statistics at the setting level (setting-specific samples are small); thus, setting-specific statistics are reported, but no hypothesis testing was performed to determine whether setting results were significantly different from one another.

Debrief Interview Feedback

Data from the debrief interviews were used to supplement the empirical evidence for feasibility of administration. Detailed notes from the debrief interviews were analyzed from a content-oriented perspective, assessing message intention, meaning, and accuracy. Common themes and key issues were identified either as those that were raised by more than one nurse describing a similar experience or were noted by one nurse but that were likely to reoccur and noticeably affect the assessment. Multiple reviewers analyzed the content of the debrief interviews to reach consensus. Comments were first identified as pertaining to either individual assessment categories or to the assessment as a whole. For categories containing multiple comments, comments were further subdivided by content area (e.g., data element, patient factors, suggestions for training). All illustrative examples provided during the debrief interview were included in the summaries. Additionally, we attempted to discern whether setting differences were evident through a stratified analysis.

Continuous Improvement Feedback

The data-collection monitoring process yielded useful information about the feasibility of cross-setting field test administration from a logistics perspective. We documented and cataloged all comments pertaining to logistical challenges and barriers, including use of the electronic tablets, gaining research nurse access to facility/agency electronic health records (EHRs), scheduling paired assessments, identifying potential patients for assessment, and completing assessments within the specified assessment window. We also noted any discrepancies in expected and actual completion rates in each setting (especially noncommunicative and discharge assessments) and noted reasons for those discrepancies. These lessons will be used in the planning, design, and implementation of the national Beta test.

Chapter Three. Overview of Sample and Data Collection

Overview of Alpha 2 Providers and Completed Assessments

Fifteen providers were recruited: four HHAs, four IRFs, three LTCHs, and four SNFs, with three providers in Denver, four in Houston, and eight in Chicago. However, one HHA did not contribute any assessments, and, therefore, the final facility/agency sample was 14 providers (see Table 3.1).

Table 3.1. Total, Average, and Range of Completed Assessments, by Assessment Type

	Target per Provider	Number per Provider Type (Average, Range)				
		HHA	IRF	LTCH	SNF	Overall
Number of providers		3	4	3	4	14
Number of paired assessments						
Communicative admission	10	27 (9.0, 3–17)	38 (9.5, 1–14)	24 (8.0, 3–11)	29 (7.3, 4–11)	118
Communicative discharge	5	6 (2.0, 0–6)	16 (4.0, 0–6)	9 (3.0, 2–4)	11 (2.8, 0–5)	42
Noncommunicative	10	2 (0.7, 0–2)	9 (2.3, 0–9)	24 (8.0, 4–11)	9 (2.3, 0–6)	44

NOTE: Excludes assessments done by only one assessor (research nurse or facility staff member only).

The overall target assessment completion goals were 120 communicative admission assessments (30 per setting type), 60 communicative discharge assessments (15 per setting type), and 120 noncommunicative assessments (30 per setting type). These goals were based on the expectation that agencies/facilities would be able to collect approximately eight to ten communicative admission and noncommunicative assessments each during the field period, and that approximately half of the patients/residents undergoing Alpha 2 admission assessments would be discharged during the field period and thus would be eligible for a discharge assessment. A total of 118 communicative admission assessments, 42 communicative discharge assessments, and 44 noncommunicative assessments were completed by both a research nurse and facility/agency staff assessor and were submitted electronically over the ten-week data-collection period.

Although the total number of paired communicative admission assessments was very close to the target (118 of 120), there was significant variation in the number of assessments contributed by each provider. Each PAC provider was asked to complete ten communicative admission assessments; providers actually completed between one and 17 assessments. We anticipated

some challenges in reaching these targets in the short field period and therefore recruited extra providers to offset the likelihood that some providers would be unable to reach the targets.

Despite recruitment of extra providers, the total numbers of paired communicative discharge assessments and noncommunicative assessments did not meet the targets (42 of 60 and 44 of 120, respectively). There was also significant variation in the number of these assessments contributed by each provider. For communicative discharge assessments, providers completed between zero and six assessments (with a goal of five). For noncommunicative assessments, providers completed between zero and 11 assessments (with a goal of ten). More noncommunicative assessments were completed in the LTCHs than in the other settings.

The number of completed assessments also differed somewhat by assessor type, particularly for the noncommunicative assessments, as shown in Table 3.2. In most cases, research nurses completed a few more communicative admission assessments and noncommunicative assessments, and the same number of discharge assessments, compared with the facility/agency staff assessors. The discrepancies in the numbers of completed assessments by research and facility/agency assessors were reportedly due to miscommunications and tracking errors. These sample sizes remain consistent for all tested data elements and thus are not repeated throughout the report. In the chapters that follow, most summary statistics (e.g., frequencies) are reported based on data from the 118 admission assessments completed by the facility/agency staff assessors. However, time to complete is reported for both assessor types, and IRR is calculated based on the paired assessments.

Table 3.2. Number of Assessments Completed, by Research Nurse and Facility Staff Assessors

	Communicative Admission Assessments	Discharge Assessments	Noncommunicative Assessments
Target	120	60	120
Research nurse	121	42	51
Facility/agency staff	118	42	45
Total paired assessments	118	42	44

Factors in Data Collection

In the sections below, we consider some of the factors that may have contributed to variability in success in reaching data-collection goals across providers and assessments.

Communicative Admission Assessments

Many providers met their goal of ten communicative admission assessments. However, several providers completed fewer than five communicative admission assessments. One HHA

did not complete any assessments because it had a large number of non–English-speaking patients, which illustrates the eventual need for the availability and possibly the testing of translated standardized data elements. Furthermore, one of the recruited IRFs completed only one assessment, reportedly because the data collection was interfering with the staff's regular responsibilities. Other issues reported by research nurses included facility/agency staff assessor vacations or unforeseen leaves of absence, providers using a single facility/agency assessor even though additional staff were trained, and fewer-than-anticipated admissions during the data-collection period. Overall, successful completion of communicative admission assessments was highly provider-specific and was dependent on the combination of the research nurse assigned to that provider, the staff assigned to collect data, provider leadership support and involvement, and such factors as timely access to the electronic medical record, staffing and availability of providers, and scheduling logistics. In addition, some providers were able to develop processes for integrating the test assessment into existing workflows to facilitate completion, while other providers were not able to dedicate extra resources to do so. Some setting-specific factors were noted: Namely, IRF and HHA patients were generally less sick, which facilitated conducting the assessment, whereas SNF and LTCH patients/residents tended to have more-severe health problems, which could have complicated successful completion of assessments.

Discharge Assessment

Completion of discharge assessments was challenging for a few reasons. SNFs and LTCHs typically had long durations of stay, which yielded fewer opportunities than anticipated to conduct discharge assessments and made the targeted goal unattainable. In addition, research nurses reported that it was difficult to facilitate the identification of discharge patients/residents for assessment. The field staff assessors were infrequently involved in the discharge date-setting and process, so the research nurses and staff had to establish contacts and relationships with case managers or discharge coordinators to help identify discharges. This was particularly challenging to do in the HHA setting, where off-site field staff were not in close contact with discharge-planning staff. In addition, providers that had large numbers of non–English-speaking patients/residents had fewer admission assessments and thus fewer discharge assessments than other providers.

Noncommunicative Assessments

There was a variety of factors contributing to the difficulty in reaching data-collection goals for the noncommunicative assessments. Many providers had few or no patients/residents who met the eligibility criteria for noncommunicative assessments. In other cases, some patients/residents were initially identified as noncommunicative, but when the assessment was initiated, the assessors realized that the patient/resident was actually communicative. Other patients/residents were truly vegetative and did not respond to stimuli and therefore were ineligible. It is possible that the providers may not have fully understood the noncommunicative

eligibility criteria. The definition of *noncommunicative* for the testing was sometimes interpreted loosely or was colored by staff members' own understanding and thus may not have been correctly applied. Additionally, there was an issue with some patients/residents or families refusing assent despite initially agreeing to participate (e.g., because a different family member was at the bedside at the initiation of the assessment). The lower-than-expected admission rates at providers also affected the number of noncommunicative patients/residents for testing. There were low rates of noncommunicative patients in the IRFs and HHAs, perhaps because of the types of patients typically cared for in these settings. Medicare requires that IRF patients receive at least three hours of therapy per day, five days per week. It might be less likely that this intensity of therapy would be prescribed for noncommunicative patients. In addition, noncommunicative patients might be less likely to be discharged to their homes than to another PAC setting.

Recommendations

Many of the challenges faced in completing the discharge and noncommunicative assessments were specific to the testing design and would not pose an issue if integrated into the current assessment workflow of HHAs, IRFs, LTCHs, and SNFs. However, while conducting further testing of these data elements (i.e., the national Beta test), implementing the following strategies might mitigate the risks for lower-than-targeted completed assessments for both discharge and noncommunicative assessments:

- Materials for the recruitment of provider participants should explicitly indicate that the field test will only include English-speaking patients/residents to avoid enrollment of providers with a preponderance of non–English-speaking patients/residents in the data-collection effort.
- To encourage patients/residents to participate (and minimize refusals), the project information sheet and accompanying assent language could be refined to be more easily understood and positively framed.
- The field staff training around obtaining assent could be enhanced to help staff introduce the project and explain the requirements to a range of different patients/residents (e.g., communicative and noncommunicative) and their family members.
- To improve provider engagement and buy-in in the project, research nurses could engage provider leadership through regular emails (e.g., a project newsletter), progress reports, or webinars on the data-collection process.
- To improve the process of identifying and completing discharge assessments, the research nurse and field staff trainings could be enhanced to focus more specifically on the discharge assessments and ways to identify planned discharges. Training could also focus on the importance of establishing a separate process for identification of these patients/residents and connecting to other provider staff who are not directly involved in the project. For example, one potential strategy is for research nurses and facility/agency staff to communicate with discharge planners and care coordinators to ensure that all discharges of included patients/residents are captured.

We also recommend strategies specific to improving the rates of noncommunicative assessments in future testing:

- To ensure that all noncommunicative patients/residents are captured, research nurses could meet with the field staff and discuss the noncommunicative eligibility criteria. This process could also be used to discuss the discharge process and changes in discharge dates.
- Field staff could be periodically reeducated on the study definition of *noncommunicative*. For patients/residents who are deemed ineligible because they do not respond at all to stimuli (e.g., persistent vegetative state, comatose), assessors could be trained to check back on them. Their status could change over the course of the data collection.
- To increase the number of eligible noncommunicative patients, Medicaid or private-pay patients/residents could be included.

Chapter Four. Results for Cognitive Status

Patients/residents in PAC settings are at risk for a number of cognitive impairments that can affect nearly every aspect of their lives.[1] As people age, changes within the brain create mild impairments in memory and information processing. Declines in cognitive function vary across individuals and can include changes in executive function, memory, and language capabilities. Conducting cognitive assessments is critically important to screen for cognitive impairment, to rate the severity of disorder, and to develop a care plan and monitor progression.[2]

Several data elements are currently being collected in the PAC assessment instruments to evaluate cognitive status (e.g., expression of ideas and wants, the Brief Interview for Mental Status [BIMS], the Confusion Assessment Method [CAM]). Evidence for cross-setting feasibility is present for many of these assessments, and they are being considered for cross-setting standardization. Our information-gathering phase identified a significant gap in cognitive status assessment using the current data elements. That gap is in executive functioning (including cognitive skills for activities of daily living) and functional performance assessment. Therefore, we identified candidate data elements for testing in Alpha 2 to potentially fill these gaps.

Description of Items

The Cognitive Status items for testing in Alpha 2, which can be found in Appendix B, Module A, include the Developing Outpatient Therapy Payment Alternatives (DOTPA) CARE tool and the Performance Assessment of Self-Care Skills (PASS) Medication Management Task.

[1] Rebecca G. Logsdon, Laura E. Gibbons, Susan M. McCurry, and Linda Teri, "Assessing Quality of Life in Older Adults with Cognitive Impairment," *Psychosomatic Medicine*, Vol. 64, No. 3, 2002, pp. 510–519; Susan E. Campbell, D. Gwyn Seymour, William R. Primrose, Joanna E. Lynch, Edmund Dunstan, Mireia Espallargues, Giovanni Lamura, Peter Lawson, Ian Philp, Elizabeth Mestheneos, Barbara Polityńska, and Ismo Raiha, "A Multi-Centre European Study of Factors Affecting the Discharge Destination of Older People Admitted to Hospital: Analysis of In-Hospital Data from the ACME*plus* Project," *Age and Ageing*, Vol. 34, No. 5, 2005, pp. 467–475; Raphael J. Heruti, Ayala Lusky, Rachel Dankner, Haim Ring, Mark Dolgopiat, Vita Barell, Shalom Levenkrohn, and Abraham Adunsky, "Rehabilitation Outcome of Elderly Patients After a First Stroke: Effect of Cognitive Status at Admission on the Functional Outcome," *Archives of Physical Medicine and Rehabilitation*, Vol. 83, No. 6, 2002, pp. 742–749.

[2] Valerie T. Cotter, "Alzheimer's Disease: Issues and Challenges in Primary Care," *Nursing Clinics of North America*, Vol. 41, No. 1, 2006, pp. 83–93; Maud J. L. Graff, Myrra J. M. Vernooij-Dassen, Marjolein Thijssen, Joost Dekker, Willibrord H. L. Hoefnagels, and Marcel G. M. Olde Rikkert, "Community Based Occupational Therapy for Patients with Dementia and Their Care Givers: Randomised Controlled Trial," *BMJ*, Vol. 333, No. 7580, 2006, p. 1196.

The DOTPA CARE Tool Assessment

The DOTPA data elements assess cognitive function in all patients/residents to allow for a broad assessment of multiple cognitive components over time. The DOTPA data elements assessed in Alpha 2 pertain to memory, attention, and problem-solving and are coded based on the assessor's observation of the patient/resident. Specifically, the assessor first indicates whether the patient/resident has problems in any of these areas and, if so, answers follow-up questions to provide more detail about each area. The DOTPA data elements are being considered for cross-setting use to assess constructs not represented in current cognitive status data elements, such as functional performance.

The PASS Medication Management Task

The PASS Medication Management Task assesses the patient's/resident's ability to manage medications by asking him or her to perform tasks: finding, reading, and understanding medication directions and putting pills correctly into a pill box. The item indicates the level and type of assistance required to complete the medication management tasks. The PASS is being considered for cross-setting use because it assesses cognitive skills for activities of daily living and daily decisionmaking—a noted gap in the current PAC cognitive status assessment. The version tested in Alpha 2 was modified and shortened from its original form to decrease the burden of the overall assessment. This streamlined version omits tasks that determine the extent to which a patient/resident can open pill containers without help because these tasks were not considered relevant to the domain of cognition. The remaining four tasks included in the Alpha 2 version of PASS (finding, reading, and understanding medication directions and putting pills correctly into a pill box) relate to aspects of executive functioning and problem-solving and are hypothesized to predict those patients/residents who require additional support when they leave acute care hospitals.

Testing Objectives

As with all Alpha 2 data elements, the main testing objectives were to assess cross-setting feasibility and IRR (goals 1 and 2, as described in Chapter Two). An additional goal for these data elements was to evaluate their relationship with the BIMS and the extent to which the tested data elements may be filling a gap in cognitive assessment (goal 3). We included this third goal because the BIMS is a strong candidate for cross-setting standardization. Thus, one requirement for consideration of any additional cognitive status data element is that it has the potential to contribute information beyond what is provided by the BIMS. For the DOTPA, we looked at the overall sample cross-tabulation of DOTPA impairment designations with BIMS impairment levels (intact, moderately impaired, severely impaired) and calculated the chi-square difference test. Although we would expect the impairment categorizations from the two assessments to be similar (significant chi-square), if the DOTPA is adding information about cognitive status above

and beyond the BIMS, we should see some discrepancy in the impairment designations from the two instruments. To accomplish goal 3 for the PASS, we examined the correlation between the overall composite BIMS score and the PASS Total Independence Score. Moderate correlations (i.e., between 0.5 and 0.6) would indicate that the PASS is assessing a related but distinct construct. We also examined the PASS Total Independence Score according to BIMS impairment levels (intact, moderately impaired, severely impaired). We would expect PASS scores to decrease with increasing BIMS impairment categories. Finally, qualitative data obtained during the debriefing interviews were used to evaluate difficulties encountered in administering any of the cognitive items.

Results

Tables 4.1 and 4.2 show the percentage of responses for each cognitive assessment item (for DOTPA and PASS, respectively), both overall and by setting. Responses are also categorized by the type of assessor (i.e., research nurse or facility/agency staff). For descriptive purposes, the remainder of this section summarizes results based on facility/agency staff ratings only. For both the DOTPA and PASS, responses varied across settings. Results for the first DOTPA item (A5a in Table 4.1), which assesses overall problems with cognition, showed that about one-half of all patients/residents (47 percent) had some problems with cognition, but this ranged from 35 percent of patients in LTCHs to 63 percent of patients in IRFs. The remainder of the DOTPA items were completed only for those patients identified in A5a as having some problems. As can be seen, the majority of the patients/residents who had problems with cognition were rated as mildly impaired (59 percent), followed by a smaller percentage who were rated as moderately impaired (32 percent), and an even smaller percentage who were rated as severely impaired (9 percent). Patients/residents identified as having cognitive problems were further assessed on their ability to complete simple and complex problems, recall basic and complex information, and complete simple and complex activities, all with and without assistance. In assessments conducted by facility/agency staff, patients with cognitive problems in the HHA setting tended to be less impaired, showing (with only a few exceptions) a relatively higher ability to complete simple and complex problems, recall basic and complex information, and complete simple and complex activities than patients/residents with cognitive problems in other settings. The distributions of responses to these items among patients/residents identified as having cognitive problems by the facility/agency staff were similar across the other three settings (IRFs, LTCHs, SNFs).

Results for PASS in Table 4.2 show the percentage of the sample with non-missing subscale scores who were given physical, continuous verbal/visual, and occasional verbal/visual assistance for each of the four tasks. The percentage of patients/residents who required no assistance to complete a given task was generally higher in HHA and IRF settings than in other settings. In addition, items that require the patient/resident to indicate on a calendar the next time

pills should be taken (items A4a and A4d) tended to be easier to complete with no assistance compared with the items that require the patient/resident to distribute pills correctly from the first (or second) bottle (A4c and A4f, respectively). Missing data tended to be high for PASS subscale scores, especially for assessments completed by facility/agency nurses (about 36 percent had missing subscale scores, compared with about 21 percent for research nurses). The Independence Mean Score (which is a mean of the four subscale scores and ranges from 0 to 3, with a 3 indicating better functioning) was lowest among LTCH and SNF patients/residents, indicating that more assistance is needed with activities of daily living in these patients/residents.

Table 4.1. Distribution of Responses to DOTPA Items, by Assessors and PAC Settings

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Problems with cognition? (A5a)										
No (percentage)	41	56	43	37	65	65	60	62	51	53
Yes (percentage)	59	44	57	63	35	35	40	38	49	47
Missing (number)	0	0	1	0	1	1	2	3	4	4
Continuing to follow-up questions (number)	16	12	21	24	8	8	12	10	57	54
Description of problem (A5b)										
Mildly impaired (percentage)	56	58	67	63	50	50	50	60	58	59
Moderately impaired (percentage)	31	25	33	25	38	50	40	40	35	32
Severely impaired (percentage)	13	17	0	12	13	0	10	0	7	9
Missing (number)	0	0	1	0	1	1	4	3	6	4
Simple problems without assistance (A5c)										
Never or rarely (percentage)	13	8	24	13	0	0	9	10	14	9
Sometimes (percentage)	25	17	24	29	38	38	45	40	30	30
Usually (percentage)	44	25	43	33	50	38	18	30	39	31
Always (percentage)	19	50	10	25	13	25	27	20	16	30
Missing (number)	0	0	1	0	1	1	3	3	5	4
Simple problems with assistance (A5d)										
Never or rarely (percentage)	0	0	5	4	0	0	0	0	2	2

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Sometimes (percentage)	0	17	5	4	38	50	18	11	12	15
Usually (percentage)	14	8	26	21	25	25	36	56	25	25
Always (percentage)	86	75	63	71	38	25	45	33	62	58
Missing (number)	2	0	3	0	1	1	3	4	9	5
Complex problems without assistance (A5e)										
Never or rarely (percentage)	44	42	45	25	57	57	70	44	51	37
Sometimes (percentage)	19	8	30	58	0	14	20	44	21	38
Usually (percentage)	19	25	25	17	43	29	10	11	23	19
Always (percentage)	19	25	0	0	0	0	0	0	6	6
Missing (number)	0	0	2	0	2	1	4	3	8	4
Complex problems with assistance (A5f)										
Never or rarely (percentage)	14	33	16	4	57	57	60	25	30	22
Sometimes (percentage)	29	25	21	21	14	0	10	38	20	22
Usually (percentage)	14	8	37	46	0	0	20	25	22	27
Always (percentage)	43	33	26	29	29	43	10	13	28	29
Missing (number)	2	0	3	0	2	2	4	5	11	7
Basic information without assistance (A5g)										
Never or rarely (percentage)	20	17	14	13	0	0	0	0	11	9
Sometimes (percentage)	13	8	29	17	28	50	9	30	22	22
Usually (percentage)	33	8	29	33	38	50	64	50	38	33
Always (percentage)	33	67	29	38	25	0	27	20	29	35
Missing (number)	1	0	1	0	1	1	3	3	6	4
Basic information with assistance (A5h)										
Never or rarely (percentage)	0	8	5	4	0	0	0	0	2	4
Sometimes (percentage)	0	8	10	8	25	25	9	11	9	11
Usually (percentage)	21	8	40	21	63	50	9	22	32	23
Always (percentage)	79	75	45	67	13	25	82	67	57	62

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Missing (number)	2	0	2	0	1	1	3	4	8	5
Complex information without assistance (A5i)										
Never or rarely (percentage)	27	42	24	25	50	25	40	30	31	30
Sometimes (percentage)	33	0	38	46	25	38	40	20	35	30
Usually (percentage)	20	17	33	29	25	38	20	50	26	31
Always (percentage)	20	42	5	0	0	0	0	0	7	9
Missing (number)	1	0	1	0	1	1	4	3	7	4
Complex information with assistance (A5j)										
Never or rarely (percentage)	7	33	5	0	50	0	10	22	13	11
Sometimes (percentage)	0	8	15	25	25	50	40	22	17	25
Usually (percentage)	36	8	30	38	13	38	20	0	27	25
Always (percentage)	57	50	50	38	13	13	30	56	42	40
Missing (number)	2	0	2	0	1	1	4	4	9	5
Simple activities without assistance (A5k)										
Never or rarely (percentage)	13	8	14	8	13	13	9	10	13	9
Sometimes (percentage)	40	25	29	29	50	50	27	20	35	30
Usually (percentage)	27	8	43	25	25	13	27	30	33	20
Always (percentage)	20	58	14	38	13	25	36	40	20	41
Missing (number)	1	0	1	0	1	1	3	3	6	4
Simple activities with assistance (A5l)										
Never or rarely (percentage)	0	8	5	4	0	0	0	11	2	6
Sometimes (percentage)	0	8	0	4	25	63	9	11	6	15
Usually (percentage)	7	8	20	29	50	0	9	11	19	17
Always (percentage)	93	75	75	63	25	38	82	67	74	62
Missing (number)	2	0	2	0	1	1	3	4	8	5
Complex activities without assistance (A5m)										
Never or rarely (percentage)	47	42	33	29	71	63	44	33	44	38
Sometimes (percentage)	33	17	29	46	0	0	33	22	27	28

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Usually (percentage)	7	25	29	25	14	25	11	44	17	28
Always (percentage)	13	17	10	0	14	13	11	0	12	6
Missing (number)	1	0	1	0	2	1	5	4	9	5
Complex activities with assistance (A5n)										
Never or rarely (percentage)	8	25	10	8	71	50	30	20	22	20
Sometimes (percentage)	15	17	35	21	14	13	30	30	26	20
Usually (percentage)	46	17	25	33	0	13	20	20	26	24
Always (percentage)	31	42	30	38	14	25	20	30	26	35
Missing (number)	3	0	2	0	2	1	4	3	11	4

NOTES: R = research nurse. F = facility/agency staff. Percentage rows tabulate responses to each item across all possible answer categories; responses may not sum to 100 percent due to rounding. Percentage rows show the number of times each item was unknown, missing, or similar. Responses tabulated in number rows are not part of the denominator for calculating percentages.

Table 4.2. Distribution of Responses to PASS Items, by Assessors and PAC Setting

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Reports time first medication (A4a)										
Physical (percentage)	17	0	11	8	18	23	20	8	16	9
Continuous verbal/visual (percentage)	4	14	4	8	9	8	10	0	6	8
Occasional verbal/visual (percentage)	30	36	25	18	18	15	20	42	24	25
No assists (percentage)	48	50	61	64	55	54	50	50	54	57
Not attempted (number)	1	0	0	0	0	0	1	1	2	1
Missing (number)	3	13	10	2	2	11	11	16	26	42
Distributes pills correctly first bottle (A4c)										
Physical (percentage)	17	21	21	19	14	7	21	23	18	18
Continuous verbal/visual (percentage)	22	7	4	24	27	40	37	31	21	25
Occasional verbal/visual (percentage)	35	36	46	35	41	40	26	38	38	37

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
No assists (percentage)	26	29	29	22	18	13	16	8	23	19
Not attempted (number)	1	0	0	0	0	0	1	1	2	1
Missing (number)	3	13	10	1	2	9	12	15	27	38
Reports time second medication (A4d)										
Physical (percentage)	9	8	11	11	14	15	15	8	12	11
Continuous verbal/visual (percentage)	0	8	7	8	18	15	0	8	6	9
Occasional verbal/visual (percentage)	17	31	11	22	14	23	15	38	14	27
No assists (percentage)	74	54	71	58	55	46	70	46	68	53
Not attempted (number)	1	0	0	0	0	0	1	1	2	1
Missing (number)	3	14	10	2	2	11	11	15	26	42
Distributes pills correctly second bottle (A4f)										
Physical (percentage)	17	7	21	19	14	13	26	21	20	16
Continuous verbal/visual (percentage)	13	21	11	11	36	40	16	21	18	20
Occasional verbal/visual (percentage)	22	21	14	22	18	20	21	14	18	20
No assists (percentage)	48	43	54	49	32	27	37	43	43	43
Not attempted (number)	1	0	0	0	0	0	1	1	2	1
Missing (number)	3	13	10	1	2	9	12	14	27	37
Independence Mean Score (SD)	2.08 (0.90)	2.29 (0.78)	2.15 (0.98)	2.11 (0.89)	1.88 (0.92)	1.85 (0.91)	1.81 (0.95)	1.90 (0.81)	2.00 (0.93)	2.06 (0.86)
Missing (number)	4	14	10	3	2	11	14	17	30	45

NOTES: R = research nurse. F = facility/agency staff. SD = standard deviation. Entries show the percentage of the sample with non-missing subscale scores who were given physical, continuous verbal/visual, and occasional verbal/visual scores for each of the four tasks.

Feasibility

Table 4.3 shows the time, on average, to complete each set of cognitive items (DOTPA and PASS). On average, completion of the DOTPA took five minutes, but the assessment took the longest for research nurses assessing LTCH patients/residents, with an average completion time of approximately seven minutes. On average, PASS tasks were completed in six to seven minutes, but assessments tended to take longer for patients/residents in LTCHs.

Table 4.3. Mean Time Spent, in Minutes, Completing Cognitive Function Items, by PAC Setting

	HHA Mean (SD)	IRF Mean (SD)	LTCH Mean (SD)	SNF Mean (SD)	Overall Mean (SD)
DOTPA					
By research nurse	3.42 (2.98)	4.51 (3.93)	7.04 (5.84)	4.80 (5.79)	4.84 (4.82)
By facility/agency staff	3.81 (3.96)	5.84 (7.07)	5.37 (5.24)	3.45 (3.10)	4.75 (5.41)
PASS Medication Management					
By research nurse	5.30 (3.23)	5.54 (3.01)	6.45 (3.85)	6.23 (2.70)	5.84 (3.15)
By facility/agency staff	6.81 (5.13)	6.65 (5.20)	7.87 (3.35)	6.12 (2.64)	6.82 (4.36)

Interrater Reliability

Tables 4.4 and 4.5 show the results of interrater agreement analysis. Agreement for the DOTPA data elements tended to be moderate overall, but varied by item. Agreement was highest for overall ratings of the presence of problems with cognition (A5a) and a general description of the problem (A5b), and lower for more-detailed assessments of simple and complex problems with and without assistance. Setting-specific IRRs were not computed for items A5b–A5n because of insufficient setting-level sample sizes (these items were completed only among the subset of patients/residents identified in A5a as having cognitive problems). IRR of PASS scores tended to be high overall and, at the setting level, lowest in HHAs.

Table 4.4 IRR of DOTPA Items

	HHA	IRF	LTCH	SNF	Overall
Problems with cognition? (A5a)	0.71	0.67	0.62	0.75	0.70
Description of problem (A5b)	—	—	—	—	0.81
Simple problems without assistance (A5c)	—	—	—	—	0.67
Simple problems with assistance (A5d)	—	—	—	—	0.54
Complex problems without assistance (A5e)	—	—	—	—	0.54
Complex problems with assistance (A5f)	—	—	—	—	0.56
Basic information without assistance (A5g)	—	—	—	—	0.66

	HHA	IRF	LTCH	SNF	Overall
Basic information with assistance (A5h)	—	—	—	—	0.39
Complex information without assistance (A5i)	—	—	—	—	0.66
Complex information with assistance (A5j)	—	—	—	—	0.43
Simple activities without assistance (A5k)	—	—	—	—	0.43
Simple activities with assistance (A5l)	—	—	—	—	0.34
Complex activities without assistance (A5m)	—	—	—	—	0.77
Complex activities with assistance (A5n)	—	—	—	—	0.60

NOTES: Number of paired observations for A5b–A5n is 54. Setting-level IRR was not computed for A5b–A5n because of insufficient setting-level sample sizes. IRR for all items except A5a were assessed by Pearson correlation. IRR for A5a was assessed by Cohen's kappa.

Table 4.5. IRR of PASS Items

	HHA	IRF	LTCH	SNF	Overall
Reports time first medication (A4a)	0.87	0.92	0.82	0.85	0.86
Distributes pills correctly first medication (A4c)	0.51	0.94	0.88	0.64	0.78
Reports time second medication (A4d)	0.77	0.92	0.93	0.86	0.87
Distributes pills correctly second medication (A4f)	0.52	0.96	0.83	0.85	0.83
Independence Mean Score	0.71	0.98	0.93	0.89	0.92

NOTE: IRR was assessed by Pearson correlation.

Relationship with BIMS Scores

Table 4.6 provides cross-tabulations between DOTPA and BIMS impairment categories. The overall chi-square was significant, $\chi^2(6) = 50.96$, $p < 0.05$, indicating that the two assessments are similar to one another. Interpretation of these results is challenging because of the different numbers of classification groups for the two instruments. Nonetheless, the table reveals some interesting areas of agreement and disagreement. Although the majority of patients/residents categorized as "intact" by the BIMS were also coded as having no problems by the DOTPA, 29 intact BIMS patients/residents were coded as having a cognitive problem according to the DOTPA. These observed patterns in impairment categorization suggest that the DOTPA is contributing additional information about cognitive status above and beyond the BIMS.

Table 4.7 shows the means and SDs for PASS Independence Mean Scores by BIMS categories. Results reveal that Independence Mean Scores decreased as cognitive impairment severity increased, as indicated by the BIMS, lending evidence for the validity of the PASS as an assessment of cognitive status. Additionally, correlations were computed between Independence Mean Scores and BIMS composite scores and were found to be moderately positively correlated such that higher Independence Mean Scores were associated with higher BIMS composite scores ($r = 0.6$, $N = 73$) overall, with some variation across settings ($r = 0.34$–0.93, $N = 12$–35). The

range of correlations suggests that the PASS contributes unique information above and beyond the BIMS.

Table 4.6. Frequency Cross-Tabulation of DOTPA and BIMS Impairment Categories

	BIMS			
DOTPA Category	**Intact**	**Moderately Impaired**	**Severely Impaired**	**Total**
No impairment	56	3*	0*	59
Mildly impaired	21	10*	1	32
Moderately impaired	7	5	5*	17
Severely impaired	1	1	3*	5
Total	85	19	9	113

NOTES: Overall $\chi^2_{(6)} = 50.96$, $p < 0.05$. Significant contributions to overall chi-square are denoted with an asterisk.

Table 4.7. Mean and Standard Deviation of PASS Independence Score, by BIMS Impairment Category

BIMS Category	**Number**	**Mean**	**SD**
Intact	54	2.33	0.68
Moderately impaired	13	1.60	0.81
Severely impaired	6	0.62	0.63

During the debrief interviews, facility/agency staff and research nurses generally found both DOTPA and PASS straightforward to administer. They also noted some challenges, however, suggesting that the DOTPA might not be relevant to patients/residents in SNFs and LTCHs specifically and also commented that the order of questions in the DOTPA on whether assistance was needed should be switched. The assessors viewed the PASS as difficult to administer in LTCH settings and with physically limited patients/residents and those who cannot sit up in bed because of the upper-extremity mobility that is required to complete PASS tasks (e.g., laying pills out on a paper calendar).

Summary of Findings

Feasibility/ease of use: Missing data were higher than expected for the PASS subscales, mainly because raters forgot to return to the section later to complete the final ratings. This is only an issue for future testing and could possibly be addressed through better training and formatting of the electronic input device. The time to administer the instruments was not trivial and may be considered overly burdensome, especially in LTCHs. Feedback from assessors

suggested that the DOTPA and PASS were generally straightforward to administer, but administration could pose some challenges, especially in the LTCH setting, where patient complexity could lead to increased time in performing the PASS and where the DOTPA may not be as relevant.

Interrater reliability: IRR varied considerably across cognitive status items, especially for the DOTPA, with some items showing excellent reliability (e.g., 0.81) and others displaying unacceptably low reliability (e.g., 0.34). IRR was high overall (0.78–0.92) for the PASS scores.

Comparison with BIMS scores: While broad agreement with BIMS categorization of patients/residents provides evidence of the validity of the DOTPA assessment, discrepancies in impairment designation suggest that additional information about cognitive status was obtained by the DOTPA above and beyond the BIMS. Correlations between the BIMS and the PASS Independence Mean Scores were moderate overall, with some setting variability, suggesting that additional information was provided by the PASS elements than with content assessed by the BIMS. Additionally, Independence Mean Scores decreased as BIMS cognitive impairment severity increased, lending evidence to the validity of the PASS.

Recommendations

Results for the DOTPA and PASS are mixed. Although they appear to contribute unique information about cognitive status beyond the BIMS, Alpha 2 administration revealed that both the DOTPA and PASS data elements had potential problems with both feasibility and reliability. In terms of feasibility, both instruments took considerable time to complete, especially in LTCHs, and were identified as potentially problematic to administer in LTCHs and SNFs. Furthermore, the PASS coding was complex and resulted in significant amounts of missing data. The PASS showed excellent IRR, but the reliability of the DOTPA was mixed. Although many of these issues could be addressed with improvements in training and guidance, the limitations should be taken into account when considering data elements for cross-setting standardization.

Chapter Five. Results for Behavioral Signs and Symptoms

Behavioral disturbances—a patient's or resident's disruptive or dangerous physical or verbal behaviors directed either at themself or at caregivers, often signaling distress or unmet or unrecognized needs—strain the time and resources of PAC providers, disrupt care, and result in poorer patient outcomes. Patients/residents with these behaviors may require more case management time, may have poorer quality of life and interpersonal relationships, and may be at risk for injury, isolation, and inactivity. These symptoms can also disrupt the institutional or home environment and affect the safety and privacy of other patients/residents and caregivers. Exposure to aggressive behaviors can also have a negative effect on staff job satisfaction. Assessment and documentation of behavioral disturbances can help inform care planning, staffing, interventions, and patient transitions.

Behavior disturbances are not currently assessed in the IRF-PAI or LCDS, but this content is included in both OASIS and the MDS. The behavioral data element tested in Alpha 2 was derived from items in the MDS 3.0 and received strong support from the TEP, where there was general agreement that behavioral assessment was important and a strong candidate for cross-setting standardization.

Description of Items

The Behavioral Signs and Symptoms data element (referred to as Behavior items) can be found in Appendix B, Module B. The Behavior items in the Alpha 2 feasibility test first assess the presence and frequency of behavioral symptoms over the past seven days. If symptoms are exhibited, follow-up questions assess the effect of behavioral symptoms on the patient/resident and on others, including risk for physical injury, interference with patient/resident care, interference with the patient's/resident's participation in activities, intrusion on the privacy of others, or disruption of the delivery of care or living environment of others. The Behavior items also include assessment of the presence and frequency of rejection of care that is not consistent with the patient's/resident's preferences or goals.

Testing Objectives

As with all Alpha 2 data elements, the main testing objectives were to assess cross-setting feasibility and IRR (goals 1 and 2, as described in Chapter Two). In addition to analysis of the assessment data, qualitative data obtained during the debriefing interviews were used to evaluate difficulties encountered in administering any of the Behavior items.

Results

Table 5.1 shows the percentage of responses for the Behavior items, overall and by setting. Responses are also categorized by the type of assessor (i.e., research nurse or facility/agency staff). For descriptive purposes, the remainder of this section summarizes results based on facility/agency staff ratings only. Across all settings, few Alpha 2 patients/residents exhibited physical, verbal, or other behavioral symptoms. Facility/agency staff documented no patients/residents exhibiting physical behaviors, only seven of 118 patients exhibiting verbal behaviors, and four of 118 exhibiting other behavioral symptoms. Although these few exhibited behaviors occurred among patients/residents in all four settings, the majority were exhibited among HHA patients (six instances), followed by SNF residents (three instances). Only one patient in each of the IRF and LTCH settings exhibited any behavioral symptoms. Responses to follow-up questions indicate that less than one-half of these 11 expressions of verbal or other behavioral symptoms had an effect on the patient/resident (B1e–g) or on others (B1h–j). Responses to the rejection-of-care item showed a similar pattern. Very few patients/residents exhibited these behaviors. Unlike the previous behaviors, rejection of care seemed to occur more often among IRF and LTCH patients relative to HHA patients and SNF residents (there were no rejection-of-care incidents among SNF residents).

Table 5.1. Distribution of Responses to Behavior Items, by Assessors and PAC Setting

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Presence of behavioral symptoms										
Physical behavioral symptoms directed toward others (B1a)										
Behavior not exhibited (percentage)	96	100	100	100	96	100	100	100	98	100
Behavior of this type occurred 1–3 days (percentage)	4	0	0	0	4	0	0	0	2	0
Behavior of this type occurred 4–6 days (percentage)	0	0	0	0	0	0	0	0	0	0
Behavior of this type occurred daily (percentage)	0	0	0	0	0	0	0	0	0	0
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	2	1	1	1	1	2	2	4	6

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Verbal behavioral symptoms directed toward others (B1b)										
Behavior not exhibited (percentage)	84	84	100	97	100	100	90	93	93	94
Behavior of this type occurred 1–3 days (percentage)	12	8	0	0	0	0	7	4	5	3
Behavior of this type occurred 4–6 days (percentage)	4	8	0	0	0	0	0	0	1	2
Behavior of this type occurred daily (percentage)	0	0	0	3	0	0	3	4	1	2
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	2	2	9	0	1	1	2	2	14	5
Other behavioral symptoms not directed toward others (B1c)										
Behavior not exhibited (percentage)	93	92	100	100	100	95	100	96	98	96
Behavior of this type occurred 1–3 days (percentage)	7	8	0	0	0	5	0	4	2	4
Behavior of this type occurred 4–6 days (percentage)	0	0	0	0	0	0	0	0	0	0
Behavior of this type occurred daily (percentage)	0	0	0	0	0	0	0	0	0	0
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	2	1	0	1	2	2	2	4	6
Follow-up: Impact on patient/resident (percentages based on <10 patients/residents exhibiting behaviors in B1a–B1c)										
Put patient/resident at significant risk for physical illness or injury (B1e)										
Yes (percentage)	50	25	N/A	0	0	0	0	0	25	13
Skipped correctly (number)	23	23	38	37	23	23	29	27	113	110
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	0	0	0	0	0	0

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Significantly interfere with the patient's/resident's care (B1f)										
Yes (percentage)	25	25	N/A	0	0	0	33	0	25	13
Skipped correctly (number)	23	23	38	37	23	23	29	27	113	110
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	0	0	0	0	0	0
Significantly interfere with the patient's/resident's participation in activities or social interaction (B1g)										
Yes (percentage)	50	25	N/A	0	0	100	0	50	25	34
Skipped correctly (number)	23	23	38	37	23	23	29	27	113	110
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Not applicable (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	0	0	0	0	0	7
Follow-up: Impact on others (percentages based on <10 patients/residents exhibiting behaviors in B1a–B1c)										
Put others at significant risk for physical injury (B1h)										
Yes (percentage)	0	0	N/A	0	0	0	0	0	0	0
Skipped correctly (number)	23	23	38	37	23	23	29	27	113	110
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	0	0	0	0	0	0
Significantly intrude on the privacy or activity of others (B1i)										
Yes (percentage)	0	0	N/A	0	0	0	0	0	0	0
Skipped correctly (number)	23	23	38	37	23	23	29	27	113	110
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	0	0	0	0	0	0
Significantly disrupt the delivery of care or living environment of others (B1j)										
Yes (percentage)	0	0	N/A	0	0	0	33	50	13	13
Skipped correctly (number)	23	23	38	37	23	23	29	27	113	110

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	0	0	0	0	0	0
Frequency of rejection of care										
Rejection of care (B1k)										
Behavior not exhibited (percentage)	85	96	89	89	91	91	97	100	91	94
Behavior of this type occurred 1–3 days (percentage)	15	4	11	11	9	9	3	0	9	6
Behavior of this type occurred 4–6 days (percentage)	0	0	0	0	0	0	0	0	0	0
Behavior of this type occurred daily (percentage)	0	0	0	0	0	0	0	0	0	0
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	2	1	0	1	1	2	2	4	5

NOTES: R = research nurse. F = facility/agency staff. SD = standard deviation. Skipped answers are appropriately missing based on previous answers and programmed skip patterns. Percentage rows tabulate responses to each item across all possible answer categories listed below the item; responses may not sum to 100 percent due to rounding. Number rows show the number of times each item was unknown, missing, or similar. Responses tabulated in the number rows were not included as part of the denominator for calculating percentages. Cells with N/A indicate no endorsement of any of the response answers (i.e., non-missing categories).

Feasibility

Table 5.2 shows the time, on average, to complete the Behavior items. Overall, both research nurses and facility/agency staff took less than three minutes to complete these items. There were some differences across settings. Facility/agency staff in SNFs took less than two minutes on average.

Table 5.2. Mean Time Spent, in Minutes, Completing Behavior Items by PAC Setting

	HHA Mean (SD)	IRF Mean (SD)	LTCH Mean (SD)	SNF Mean (SD)	Overall Mean (SD)
By research nurse	1.88 (1.21)	2.49 (1.97)	3.43 (2.13)	3.64 (2.53)	2.82 (2.11)
By facility/agency staff	3.08 (2.46)	3.16 (4.39)	3.50 (2.67)	1.89 (1.19)	2.90 (3.12)

Interrater Reliability

Interrater agreement was difficult to evaluate for the Behavior items because of the very small number of behaviors that were exhibited. There were not enough behavioral symptoms exhibited to calculate IRR by setting, and there were so few patients/residents for whom the effect on the patient/resident and the effect on other items were completed that we were unable to calculate IRR for items B1d–B1j (these items were appropriately skipped for patients/residents when physical, verbal, and other behavioral symptoms were not exhibited). Similarly, IRR for B1a was not defined because there was no variability in responses among facility/agency staff ratings: All facility/agency staff ratings for this item were "behavior not exhibited." However, we were able to calculate overall IRR for items B1b, B1c, and B1k, which showed moderate to good reliability (kappa = 0.77, 0.66, and 0.60, respectively).

Feedback from Assessors

Facility/agency staff and research nurses used a variety of sources—staff, caregivers, medical record—to complete the Behavior items. Several facility/agency staff members reported that staff were helpful in answering the Behavior items, but one noted that staff have limited exposure to the behavior of patients/residents since they change at each shift. One facility/agency staff assessor in an LTCH reported that his or her patients exhibit apathy or are discouraged, but tend not to exhibit many behavioral disturbances. Assessors also reported that it was difficult to use the full seven-day assessment window embedded in the Behavior item response options (behavior occurred one to three days, four to six days, or daily) given the Alpha 2 study design. All Alpha 2 admission assessments were completed on day three or day four of admission and assessors had difficulty obtaining preadmission information regarding behavior disturbances. Thus, they had difficulty ascertaining whether behaviors occurred prior to admission and therefore could not use the full set of response options.

Summary of Findings

Feasibility/ease of use: The behavioral signs and symptoms items took less than three minutes to complete on average. Assessors had difficulty considering the full seven-day look-back period when completing these items because Alpha 2 assessments had to be completed by day four, and assessors had difficulty obtaining preadmission information about patients/residents related to behavior.

Interrater reliability: Interrater agreement was challenging to evaluate because of small sample sizes and little variability in responses. However, IRR was moderate to good for the three items for which it could be calculated.

Recommendations

Results from Alpha 2 are mixed, and very few behavioral symptoms were exhibited among the patients/residents in this pilot test, making evaluation of psychometric performance challenging. However, the Behavior items do not appear to be overly burdensome or problematic for assessors to complete. Although the use of the seven-day assessment window was problematic in this test given its lack of correspondence with the study design requirements, this issue can be addressed in the future by matching the items' look-back period to the assessment period. In future testing, the look-back period will fall within the current stay, and, therefore, assessors will not need to rely on preadmission information. In addition, different look-back periods should be tested to determine which length is most appropriate for capturing behavioral symptoms.

Chapter Six. Results for Anxiety

Anxiety disorders are the most common lifetime mental health disorders in the United States that continue to be prevalent in older populations (ages 65 and older).[1] Anxiety is not a normal response to physical ailments or loss of independence; it is a medical illness that is unlikely to go away on its own.[2] The disorders can cause significant impairment in life function and are associated with increased clinical care needs and resource use.[3] Undetected anxiety disorders can complicate depression, pain, and disease management.[4] However, these disorders are responsive to treatment if detected.[5] Thus, psychometrically sound assessment instruments for anxiety in older adults are greatly needed.[6]

Given the prevalence and potential problems associated with anxiety symptoms, CMS is considering anxiety data elements for possible inclusion in cross-setting standardized patient assessments. At this time, anxiety-related assessment data elements are not included on the four assessment instruments. By documenting the frequency of specific indicators of anxiety, staff in PAC settings can begin to recognize anxiety indicators and consider them when developing the patient's/resident's individualized care plan, regardless of whether the patient/resident meets the criteria for a diagnosis of an anxiety disorder. This chapter describes the data elements tested in Alpha 2 that assess anxiety. The items were selected from the Patient-Reported Outcomes Measurement Information System (PROMIS) Anxiety item bank.[7]

Description of Items

The PROMIS Anxiety item bank focuses on assessing self-reported fear (fearfulness, panic), anxious misery (worry, dread), hyperarousal (tension, nervousness, restlessness), and somatic symptoms related to arousal (racing heart, dizziness). It has a total of 29 items, from which 11 were selected for testing in Alpha 2. These items were selected based on feedback from PAC stakeholders and clinical experts who rated the 29 items. The selected 11 items assess a wide

[1] Landis and Koch, 1977.

[2] Logsdon et al., 2002.

[3] Campbell et al., 2005; Heruti et al., 2002.

[4] Cotter, 2006; Graff et al., 2006.

[5] Rosalie A. Kane, "Goals of Home Care: Therapeutic, Compensatory, Either, or Both?" *Journal of Aging and Health*, Vol. 11, No. 3, 1999, pp. 299–321.

[6] Robert L. Kane and Rosalie A. Kane, "What Older People Want from Long-Term Care, and How They Can Get It," *Health Affairs*, Vol. 20, No. 6, 2001, pp. 114–127.

[7] PROMIS is a National Institutes of Health (NIH) Roadmap Initiative to develop standardized item banks to assess self-reported physical, mental, and social health.

range of anxiety symptom severity. They ask the respondent to report the frequency of symptom experience in the past seven days and are coded on a rating scale from 1 to 5, where 1 = *never*, 2 = *rarely*, 3 = *sometimes*, 4 = *often*, and 5 = *always*. Additional response codes are 7 = *patient/resident declined to respond*, and 9 = *unknown or unable to assess*.

Anxiety items tested in Alpha 2 are presented in Appendix B, Module D.

Testing Objectives

As with all Alpha 2 data elements, the main testing objectives were to assess cross-setting feasibility and IRR (goals 1 and 2, as described in Chapter Two). An additional goal for the Anxiety data elements was to determine how Anxiety scores in the PAC settings compare with general-population norms (goal 3). Achieving this goal is straightforward because the Anxiety items are a subset of the PROMIS item bank. Thus, we were able to calculate an overall scale score and convert them to a T-score metric for comparison with the published general-population norms.

Qualitative data obtained from nurse feedback during the debriefing interviews were used to evaluate feasibility of administration and identify challenges with assessment.

Results

Table 6.1 shows the percentage of responses to each Anxiety item, overall and by setting. Across all settings, responses to each item varied, indicating that assessors used the full range of response options when completing the assessment. For symptoms that are more common (e.g., "I felt worried"), most patients endorsed *sometimes*. For symptoms that are more severe (e.g., "My worries overwhelmed me"), most patients endorsed *never* or *rarely*. There was also considerable variability in the distribution of responses across settings. For example, patients from LTCH and SNF settings tended to report experiencing anxiety symptoms more frequently than those in the HHA and IRF settings.

Table 6.1. Distribution of Responses to Anxiety Items, by Assessors and PAC Setting

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
I had difficulty sleeping (D1a)										
1. Never (percentage)	19	19	16	16	5	9	10	11	10	14
2. Rarely (percentage)	15	15	16	16	5	5	7	7	7	11
3. Sometimes (percentage)	33	33	29	29	45	45	47	44	47	37
4. Often (percentage)	15	15	29	29	23	27	27	30	27	25
5. Always (percentage)	19	19	11	11	23	14	10	7	10	12
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	0	0	1	1
9. Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	1	1	2	2	3	3
I felt worried (D1b)										
1. Never (percentage)	26	26	16	16	14	14	13	11	17	17
2. Rarely (percentage)	22	22	26	24	9	9	23	26	21	21
3. Sometimes (percentage)	30	30	32	34	27	27	43	44	33	34
4. Often (percentage)	7	7	13	11	23	23	13	15	14	13
5. Always (percentage)	15	15	13	16	27	27	7	4	15	15
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	0	0	0	1
9. Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	1	1	2	2	2	3
My worries overwhelmed me (D1c)										
1. Never (percentage)	48	48	47	46	18	18	40	41	40	40
2. Rarely (percentage)	19	19	18	19	23	23	17	15	19	19
3. Sometimes (percentage)	22	26	21	22	23	23	30	33	24	26

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
4. Often (percentage)	7	4	11	11	9	9	3	4	8	7
5. Always (percentage)	4	4	3	3	27	27	10	7	9	8
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	1	0	1	1
9. Unknown or unable to assess (number)	0	0	0	1	0	0	0	0	0	1
Missing (number)	0	0	0	0	1	1	1	2	3	3
I had trouble paying attention (D1d)										
1. Never (percentage)	33	37	29	29	14	14	13	15	23	25
2. Rarely (percentage)	30	30	16	16	27	27	40	41	27	27
3. Sometimes (percentage)	30	30	37	37	32	32	27	22	32	31
4. Often (percentage)	7	4	16	16	23	23	10	11	14	13
5. Always (percentage)	0	0	3	3	5	5	10	11	4	4
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	0	0	0	1
9. Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	1	1	2	2	2	3
I felt nervous (D1e)										
1. Never (percentage)	33	33	18	18	9	9	13	15	19	19
2. Rarely (percentage)	26	22	24	24	9	9	30	30	23	22
3. Sometimes (percentage)	22	22	37	37	64	64	40	41	39	39
4. Often (percentage)	15	19	18	18	9	9	10	11	14	15
5. Always (percentage)	4	4	3	3	9	9	7	4	5	4
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	0	0	1	1
9. Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Missing (number)	0	0	0	0	1	1	2	2	3	3
I felt anxious (D1f)										
1. Never (percentage)	31	30	18	18	18	18	10	11	19	19
2. Rarely (percentage)	12	11	24	24	23	23	33	33	23	23
3. Sometimes (percentage)	38	41	45	45	41	41	43	44	42	43
4. Often (percentage)	8	7	11	11	9	9	10	11	9	10
5. Always (percentage)	12	11	3	3	9	9	3	0	6	5
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	0	0	1	1
9. Unknown or unable to assess (number)	1	0	0	0	0	0	0	0	1	0
Missing (number)	0	0	0	0	1	1	2	2	3	3
I had difficulty calming down (D1g)										
1. Never (percentage)	44	48	37	37	18	18	33	37	34	36
2. Rarely (percentage)	33	30	24	24	27	27	33	30	29	27
3. Sometimes (percentage)	11	11	26	24	32	32	23	26	23	23
4. Often (percentage)	7	7	13	13	23	23	0	0	10	11
5. Always (percentage)	4	4	0	3	0	0	10	7	3	4
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	0	0	1	1
9. Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	1	1	2	2	3	3
I had a racing or pounding heart (D1h)										
1. Never (percentage)	63	60	61	61	45	45	47	44	55	54
2. Rarely (percentage)	4	7	16	16	14	14	27	30	15	17
3. Sometimes (percentage)	33	33	18	18	36	36	20	22	26	26
4. Often (percentage)	0	0	3	3	5	5	3	4	3	3

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
5. Always (percentage)	0	0	3	3	0	0	3	0	2	1
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	0	0	1	1
9. Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	1	1	2	2	3	3
I found it hard to focus on anything other than my anxiety (D1i)										
1. Never (percentage)	48	50	49	49	14	14	27	30	36	38
2. Rarely (percentage)	30	27	24	24	41	41	40	37	33	31
3. Sometimes (percentage)	19	19	14	14	27	27	20	22	19	20
4. Often (percentage)	4	4	14	14	14	14	10	11	10	11
5. Always (percentage)	0	0	0	0	5	5	3	0	2	1
7. Patient/resident declined to respond (number)	0	0	1	0	1	1	0	0	2	1
9. Unknown or unable to assess (number)	0	0	0	1	0	0	0	0	0	1
Missing (number)	0	1	0	0	1	1	2	2	3	4
I felt like I needed help for my anxiety (D1j)										
1. Never (percentage)	65	63	55	55	27	23	53	52	52	50
2. Rarely (percentage)	12	15	18	18	18	23	20	22	17	19
3. Sometimes (percentage)	15	15	11	11	36	26	20	22	19	19
4. Often (percentage)	8	7	16	16	14	14	3	4	10	11
5. Always (percentage)	0	0	0	0	5	5	3	0	2	1
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	0	0	1	1
9. Unknown or unable to assess (number)	1	0	0	0	0	0	0	0	1	0
Missing (number)	0	0	0	0	1	1	2	2	3	3
I had sudden feelings of panic (D1k)										

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
1. Never (percentage)	67	67	45	42	36	36	47	44	49	47
2. Rarely (percentage)	22	22	29	32	23	23	37	41	28	30
3. Sometimes (percentage)	11	11	16	16	32	32	10	11	16	17
4. Often (percentage)	0	0	11	11	9	9	0	0	5	5
5. Always (percentage)	0	0	0	0	0	0	7	4	2	1
7. Patient/resident declined to respond (number)	0	0	0	0	1	1	0	0	1	1
9. Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	1	1	2	2	3	3

NOTES: Percentage rows tabulate responses to each item across all possible answer categories. Responses may not sum to 100 percent due to rounding. Responses tabulated in number rows are not part of the denominator for calculating percentages.

Feasibility

Table 6.2 shows the time, on average, to complete the set of 11 Anxiety items. Overall, the average time was almost four minutes for research nurses and almost five minutes for facility/agency staff, with some variation across settings. The assessment appeared to take longer in LTCHs than in other settings.

Table 6.2. Mean Time Spent, in Minutes, Completing Anxiety Items, by PAC Setting

	HHA Mean (SD)	IRF Mean (SD)	LTCH Mean (SD)	SNF Mean (SD)	Overall Mean (SD)
By research nurse	3.42 (1.86)	2.86 (1.38)	4.55 (2.54)	3.60 (1.73)	3.50 (1.91)
By facility/agency staff	5.22 (2.58)	4.50 (2.58)	6.37 (2.52)	3.85 (1.79)	4.84 (2.51)

Interrater Reliability

Table 6.3 shows IRR for the Anxiety items across paired observations. Overall, kappas ranged from 0.80 to 1.00, with the majority between 0.95 and 1.00, indicating near-perfect interrater agreement for almost all items. This suggests that different assessors at all facilities were able to obtain extremely similar results for these items.

Table 6.3. IRR of Anxiety Items

	HHA	IRF	LTCH	SNF	Overall
I had difficulty sleeping (D1a)	1.00	1.00	0.80	1.00	0.97
I felt worried (D1b)	0.95	0.96	1.00	1.00	0.98
My worries overwhelmed me (D1c)	0.97	1.00	1.00	1.00	0.99
I had trouble paying attention (D1d)	0.89	1.00	1.00	1.00	0.98
I felt nervous (D1e)	0.89	1.00	1.00	1.00	0.97
I felt anxious (D1f)	1.00	1.00	1.00	1.00	1.00
I had difficulty calming down (D1g)	0.97	N/A	1.00	0.97	0.97
I had a racing or pounding heart (D1h)	0.96	1.00	1.00	0.96	0.98
I found it hard to focus on anything other than my anxiety (D1i)	1.00	1.00	1.00	1.00	1.00
I felt like I needed help for my anxiety (D1j)	1.00	1.00	0.96	1.00	0.99
I had sudden feelings of panic (D1k)	1.00	0.98	1.00	1.00	0.99

NOTES: IRR was assessed by weighted kappa. Cells with N/A indicate that not all response categories were endorsed by both nurses; thus, IRR cannot be computed.

Comparison with the General Population

Table 6.4 shows Anxiety T-score means and SDs based on facility/agency staff ratings for each setting and overall. The scores are on a T-score metric where the general-population norm is a mean of 50 with an SD of ten (the conversion table is shown as Table 6.5 for reference). As can be seen in Table 6.4, patients/residents in Alpha 2 were, on average, more than half an SD above the general-population mean Anxiety scores, indicating higher overall anxiety. There is also slightly lower variability in the Alpha 2 sample relative to general-population norms, perhaps because of the relatively small sample size. At the setting level, HHA patients scored the lowest on average, whereas LTCH patients displayed the highest levels of anxiety symptoms. SNF residents and IRF patients endorsed anxiety symptoms at similar rates.

Table 6.4. Mean Anxiety T-Scores by PAC Setting, Based on Facility Staff Assessment Data

	HHA Mean (SD)	IRF Mean (SD)	LTCH Mean (SD)	SNF Mean (SD)	Overall Mean (SD)
Anxiety	53.2 (8.9)	56.0 (8.2)	59.6 (8.8)	56.2 (6.6)	56.1 (8.3)

NOTE: General-population mean (SD) = 50 (10).

Table 6.5. Raw Score to T-Score Conversion for PROMIS Anxiety 11-Item Set Collected in PAC Cross-Setting Alpha 2 Feasibility Test

Raw Score	T-Score	Raw Score	T-Score	Raw Score	T-Score	Raw Score	T-Score
11	35.0	23	54.9	35	64.6	47	74.9
12	39.7	24	55.8	36	65.4	48	75.9
13	42.1	25	56.6	37	66.3	49	77.0
14	44.3	26	57.4	38	67.1	50	78.1
15	46.0	27	58.2	39	67.9	51	79.4
16	47.6	28	59.0	40	68.7	52	80.8
17	48.9	29	59.8	41	69.6	53	82.4
18	50.1	30	60.6	42	70.4	54	84.1
19	51.2	31	61.4	43	71.3	55	85.2
20	52.2	32	62.2	44	72.2		
21	53.1	33	63.0	45	73.0		
22	54.0	34	63.8	46	74.0		

Feedback from Assessors

Both facility/agency staff and research nurses indicated that, while the Anxiety items are easy to administer, the time it takes to complete them and their repetitive content are burdensome for patients/residents. One research nurse suggested that a skip pattern would be helpful so that patients/residents without anxiety did not have to answer all the questions. In some instances, the questions appeared to induce anxiety in patients because patients seemed to start to think that they should have anxiety. In other instances, patients/residents became upset by the items. Facility/agency staff and research nurses noted that, rather than explicitly referencing the seven-day time frame, patients/residents seemed to be answering the items more generally. The lack of definitions for the terms in the items caused patients/residents to ask for assistance. Finally, patients were not always comfortable answering these items with a nurse with whom they were not familiar.

Summary of Findings

Feasibility/ease of use: Assessors' comments generally indicated that the Anxiety items were straightforward to administer. The high consistency of ratings across assessors and the few endorsements of "unknown or unable to assess" support these comments. Although assessor feedback indicates that some patients were upset when asked the questions, few patients declined to respond. The full 11-item set took approximately five minutes to administer. The extent of missing data is also very low. Assessors reported that the instructions were clear. There were concerns that the seven-day time frame could be a challenge, since it requires patients/residents to consider their experiences on one or more days prior to admission.

Interrater reliability: The Anxiety items performed well across all PAC settings in terms of IRR. Interrater agreement on Anxiety items was almost perfect (0.80 to 1.00), with most items exceeding 0.95. Little variation was observed across settings.

Comparison with the general population: Patients/residents in Alpha 2 showed higher levels of anxiety symptoms than is seen on average in the general population.

Recommendations

Results for the 11-item Anxiety assessment show that inclusion in standardized assessment could be feasible. The items were straightforward to administer, there were low rates of missing responses, and IRR was nearly perfect. However, feedback from assessors indicates that both the assessors and patients/residents found the item set overly repetitive and burdensome and suggested that it be shortened or that a skip pattern be created for those who are identified as at low risk for anxiety symptoms. Future testing of the Anxiety data elements will take these issues into consideration. Finally, alternative time frames (e.g., "in the past 3 days . . .") will be taken into consideration as we explore reliable and valid data elements for cross-setting standardization.

Chapter Seven. Results for Care Preferences

Assessment of patient preferences for care in PAC settings is crucial to informing an individualized care plan and planning for successful care transitions.[1] In addition to clinical guidelines, information about patient preferences and goals provides important direction for developing a care plan, selecting treatment options, and tailoring interventions. Understanding patient goals can also help to establish or reset both patient and provider expectations in the context of the current clinical condition. Improved understanding of patient preferences and goals through a systematic assessment process can also strengthen the patient-provider relationship and build trust.

Currently, the assessment of patient preferences in PAC is limited and not standardized; preferences for involvement in treatment and treatment decisionmaking, preferences for provider and type of care, and overall goals for health care intervention are not addressed. Alpha 1 testing demonstrated several potential successful data elements for Care Preferences; however, there were two areas, Physician Orders and Goals of Care, for which the RAND team received specific feedback that warranted additional testing prior to the Beta test period.

In this chapter, we describe the Care Preferences data elements—Physician Orders and Goals of Care—that were tested in Alpha 2, the testing objectives, and results from the Alpha 2 feasibility test.

Description of Items

The Care Preferences items for testing in Alpha 2, which can be found in Appendix B, Module G, were Physician Orders and Goals of Care.

Physician Orders

Feedback from the public comment periods and the clinical advisers recommended expanding the advance directive item tested in Alpha 1 (Health Care Agent) to include specific treatment decisions. The TEP, held in January 2017, suggested that some expansion of the items currently used in the MDS would compose a strong candidate set of data elements for standardization that aligned with the feedback. From this feedback, a set of five Physician Orders adapted from the MDS was tested in Alpha 2 as Physician Orders. The Physician Orders data

[1] Kane, 1999; Kane and Kane, 2001; Carol J. Whitlatch, Rich Piiparinen, and Lynn Friss Feinberg, "How Well Do Family Caregivers Know Their Relatives' Care Values and Preferences?" *Dementia*, Vol. 8, No. 2, 2009, pp. 223–243; J. E. Arnetz, I. Almin, K. Bergstrom, Y. Franzen, and H. Nilsson, "Active Patient Involvement in the Establishment of Physical Therapy Goals: Effects on Treatment Outcome and Quality of Care," *Advances in Physiotherapy*, Vol. 6, No. 2, 2004, pp. 50–69.

element was assessed through medical chart review. The assessor was required to document the presence of any and all of the five Physician Orders contained within the item (do not resuscitate [DNR], do not intubate [DNI], do not hospitalize, antibiotic restrictions, comfort care preferences) or document that no Physician Orders were located in the chart review.

Goals of Care

A patient interview data element reflecting the importance of goals of care to the patient was tested in Alpha 1. Although Alpha 1 results demonstrated high feasibility, the data element showed limited variation in response patterns: Essentially all patients/residents reported that having goals of care was important. In considering these results in January 2017, the TEP was not surprised by the lack of variability and also raised concerns that a discussion about specific patient/resident goals was not well suited for standard assessment. In light of these results and comments, a greater process-based (i.e., chart review) version of the Goals of Care data element was developed and tested for feasibility in Alpha 2. For this data element, assessors documented whether there was evidence in the medical chart of a goals-of-care conversation having occurred between a provider and a patient/resident, focusing on the goals of the patient/resident. A follow-up item was completed if a goals-of-care conversation was documented, indicating (with a check mark) which of four types of potential goals of care were discussed: physical, emotional, social, or intellectual/mental. Assessors also indicated "other" and made a written notation as needed.

Testing Objectives

As with all Alpha 2 data elements, the main testing objectives were to assess cross-setting feasibility and IRR (goals 1 and 2, as described in Chapter Two). In addition to analysis of the assessment data, qualitative data obtained during the debriefing interviews were used to evaluate difficulties encountered in administering the Care Preferences data elements.

Results

Table 7.1 shows the percentage of responses for each Care Preferences data element, overall and by setting. Responses are also categorized by the type of assessor (i.e., research nurse or facility/agency staff). For descriptive purposes, the remainder of this section summarizes results based on facility/agency staff ratings only. Few assessments contained evidence of Physician Orders within the medical chart. Across settings, 75 percent of assessments had no Physician Orders listed in the medical chart, ranging from 70 percent in HHA settings to 83 percent in LTCH settings. The most common Physician Order found in the medical chart was DNR, which was found in 21 percent of the assessments. This was followed by "comfort care preference(s)," which was found in 8 percent of the assessments. In contrast, documentation of a goals-of-care conversation between a provider and a patient/resident was highly prevalent across all settings.

Overall, 80 percent of assessments exhibited evidence of a goals-of-care conversation. This was most frequently found in HHA settings (89 percent), but also in SNF (79 percent), LTCH (77 percent), and IRF (76 percent) settings. When there was evidence of a goals-of-care conversation, the most common type of conversation related to physical goals (99 percent of assessments contained evidence of a physical-goal conversation). There were low levels of missing data across the Care Preferences data elements, suggesting that research nurses and facility/agency staff assessors were generally able to complete the assessment.

Table 7.1. Distributions of Responses to Care Preferences Items, by Assessors and PAC Setting

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Physician Orders (G1b)										
a. Do not resuscitate (DNR) (yes percentage)	11	15	5	24	13	17	26	25	13	21
b. Do not intubate (DNI) (yes percentage)	0	0	3	3	9	9	0	0	3	3
c. Do not hospitalize (DNH) (yes percentage)	0	0	0	0	0	0	0	0	0	0
d. Antibiotic restriction(s) (yes percentage)	7	0	0	0	0	0	0	7	2	2
e. Comfort care preference(s) (yes percentage)	7	22	3	5	13	0	3	4	6	8
z. None of the above (percentage)	89	70	92	76	87	83	74	71	86	75
Missing (number)	0	0	0	0	1	1	1	1	2	2
Goals of Care (G1C)										
Yes (percentage)	76	89	97	76	48	77	66	79	75	80
Unknown or unable to assess (number)	2	0	0	5	0	0	1	0	3	5
Missing (number)	0	0	1	0	1	2	2	1	4	3
Goals of Care (G1d, if yes, above)										
1. Physical goals (yes percentage)	100	100	100	100	73	100	95	95	95	99
2. Emotional goals (yes percentage)	0	13	8	12	55	29	26	23	16	18
3. Social goals (yes percentage)	0	38	22	8	18	18	37	36	20	25
4. Intellectual/mental goals (yes percentage)	5	0	8	24	9	12	5	9	7	11

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
5. Other (yes percentage)	0	25	0	4	27	24	5	0	5	15
Missing (number)	0	0	1	0	1	2	2	1	4	3
Skipped correctly (number)	8	3	1	13	12	5	11	6	32	27

NOTES: Percentage rows tabulate responses to each item across all possible answer categories. Number rows show the number of times each item was unknown, missing, or similar. Responses tabulated in number rows are not part of the denominator for calculating percentages.

Table 7.2 shows the average times to complete the Care Preferences data elements. Research nurses took longer to complete these data elements (around seven minutes) than facility/agency staff (around four minutes). This difference was most pronounced in IRF and SNF settings and may be explained by the varied experiences of assessors in locating information in patient/resident charts, which we discuss in more detail in the next section.

Table 7.2. Mean Time Spent, in Minutes, Completing Care Preferences Items, by PAC Setting

	HHA Mean (SD)	IRF Mean (SD)	LTCH Mean (SD)	SNF Mean (SD)	Overall Mean (SD)
By research nurse	4.74 (4.68)	8.32 (9.00)	4.18 (4.08)	8.57 (8.61)	6.77 (7.48)
By facility/agency staff	3.11 (2.22)	3.34 (2.88)	5.59 (3.11)	2.79 (1.97)	3.58 (2.74)

Table 7.3 shows interrater agreement. For the "do not hospitalize" and "antibiotic restrictions" categories of Physician Orders, no evidence was found for their presence in the medical chart in any setting, and, as such, IRR was not computed (a 0-percent entry in any cell of Table 7.1 renders kappa undefined). Similarly, "do not intubate" and "comfort care preference(s)" were found in only two of the four settings. Among the four Physician Order categories where some or all kappas were defined (including "none of the above"), IRR varied substantially. Overall kappas ranged from 0.22 to 0.66, and within-setting kappas were as low as 0.16 and as high as 1.00. The orders "do not resuscitate" and "do not intubate" both showed moderate overall kappas (0.66). IRR for Goals of Care was generally poor. Results are similar for the specific types of Goals of Care conversations (G1d), although lack of data precluded calculation of kappa for several of these types. This is especially true for the HHA setting, where IRR was unable to be calculated for any of the G1d types of goals because of a 0-percent entry by either the research nurse or the facility/agency staff assessor. Where it could be calculated, overall kappa for the G1d item was low, ranging from –0.35 to 0.49.

Table 7.3. IRR of Care Preferences Items (Number of Paired Observations = 118)

	HHA	IRF	LTCH	SNF	Overall
Physician Orders (G1b)					
a. Do not resuscitate (DNR)	0.51	0.30	0.83	1.00	0.66
b. Do not intubate (DNI)	N/A	1.00	0.45	N/A	0.66
c. Do not hospitalize (DNH)	N/A	N/A	N/A	N/A	N/A
d. Antibiotic restriction(s)	N/A	N/A	N/A	N/A	N/A
e. Comfort care preference(s)	0.16	0.65	N/A	N/A	0.22
z. None of the above	0.24	0.43	0.83	0.91	0.59
Goals of Care (G1C)					
Yes	–0.14	0.18	–0.45	0.31	–0.03
Goals of Care (G1d, if yes, above)					
1. Physical goals	N/A	N/A	N/A	1.00	0.49
2. Emotional goals	N/A	0.62	0.33	–0.24	0.20
3. Social goals	N/A	–0.15	–0.20	–0.51	–0.35
4. Intellectual/mental goals	N/A	0.41	N/A	0.00	0.35
5. Other	N/A	N/A	–0.29	0.63	0.24

NOTES: IRR was assessed by Cohen's Kappa. Cells with N/A indicate that the frequency table is too sparse to compute IRR.

Feedback from Assessors

There was limited qualitative feedback from assessors on the Physician Orders items; however, qualitative feedback received for the Goals of Care items suggested that these items were difficult to collect. Research nurses noted that medical charts did not always clearly document that goals were discussed with the patient, and feedback from both types of assessors suggested that it was not always clear what would qualify as a documented goal conversation or where to look for evidence of one. Some comments also indicated that there was variability both in the way information on goals was collected across facilities (i.e., it was documented routinely at one facility/agency but not at another) and in the types of goals that are typically discussed. Feedback and data suggested that it was more common to find evidence of a goals-of-care conversation that involved physical goals than other types of goals. One assessor indicated that goals-of-care conversations have often occurred in a previous facility/agency, such as a hospital, but that documentation may not be transferred to the current PAC facility/agency.

Summary of Findings

Feasibility/ease of use: Low levels of missing data across the assessment suggest that researchers and facility/agency staff were generally able to complete the assessment. Facility/agency staff completed the Care Preferences data elements in approximately four minutes, whereas research nurses took an average of seven minutes to complete them. However, even though the assessments were completed, feedback from assessors indicated that it was difficult to locate in the medical chart the information needed to complete these data elements.

Interrater reliability: IRR was not able to be calculated for many categories of both Physician Orders and Goals of Care data elements in multiple settings because of infrequent occurrence. For Physician Orders, IRR that could be calculated ranged widely both overall and within settings. However, IRRs were moderate for DNR and DNI orders. IRR for the Goals of Care items was generally poor.

Recommendations

Results from Alpha 2 are mixed for these particular data elements. Very few Physician Orders were documented in medical charts among the patients/residents in this pilot test (75 percent of Alpha 2 patients/residents had no documentation of any of the Physician Orders in their charts, according to facility/agency staff), and where orders were documented, agreement between facility/agency staff and research nurses was inconsistent. Nonetheless, for the relatively common order of DNR, and the very few identified orders of DNI, IRR was moderate. For the data element documenting evidence of a goals-of-care conversation and the type of goals discussed, feedback from the assessors indicated that this information was difficult to collect, and IRR for this set of items was typically low. These limitations in feasibility should be taken into account when considering these data elements for further testing and eventual cross-setting standardization.

Chapter Eight. Results for Medication Reconciliation

MR, the process of obtaining a patient's multiple medication lists and reconciling any discrepancies, can promote patient safety by reducing errors and any resulting adverse drug events. Studies have repeatedly shown that formal MR can improve quality of life and reduce morbidity and mortality.[1] MR was adopted by the Joint Commission as a National Patient Safety Goal in 2005.[2] The five steps in the Joint Commission's MR process are (1) develop a list of current medications, (2) develop a list of medications to be prescribed, (3) compare medications on the two lists, (4) make clinical decisions based on the comparisons, and (5) communicate the new list to the patient and appropriate caregivers.[3] Development of standardized items is important to assessing whether MR aids in the improvement of patient care at points of transition while reducing medication errors.

In this chapter, we describe the MR items that were developed for consideration in cross-setting standardized assessment and tested in Alpha 2, the testing objectives, and results. We also provide a summary of findings and a set of recommendations.

Description of Items

The MR items tested in Alpha 2, which can be found in Appendix B, Module F, were developed to assess whether and how MR was conducted. The goal was to create a standardized set of items that assesses the MR process with clear definitions of each step to better explain processes for providers aiming to improve care and ease care transitions. An initial version of the MR items was tested during the Alpha 1 test and revised in light of the test findings and feedback from CMS, the January 2017 TEP, and a panel of federal subject-matter experts. The standardized items tested in Alpha 2 do not involve the assessor conducting MR. Rather, the items involve using information sources (e.g., nurse notes, medication administration records, discharge summaries, patient's medication lists) to identify which medications (including dose, route, and frequency) patients/residents are taking, determine whether there are any documented indications and discrepancies, identify whether there was reconciliation of discrepancies, and

[1] Thomas Delate, Elizabeth A. Chester, Troy W. Stubbings, and Carol A. Barnes, "Clinical Outcomes of a Home-Based Medication Reconciliation Program After Discharge from a Skilled Nursing Facility," *Pharmacotherapy*, Vol. 28, No. 4, 2008, pp. 444–452; Yuhua Bao, Huibo Shao, Tara F. Bishop, Bruce R. Schackman, and Martha L. Bruce, "Inappropriate Medication in a National Sample of U.S. Elderly Patients Receiving Home Health Care," *Journal of General Internal Medicine*, Vol. 27, No. 3, 2012, pp. 304–310.

[2] The Joint Commission, *Comprehensive Accreditation Manual for Hospitals*, Chicago, Ill., 2015.

[3] The Joint Commission, 2015; Kenneth S. Boockvar, Heather Carlson LaCorte, Vincent Giambanco, Bella Fridman, and Albert Siu, "Medication Reconciliation for Reducing Drug-Discrepancy Adverse Events," *American Journal of Geriatric Pharmacotherapy*, Vol. 4, No. 3, 2006, pp. 236–243.

determine whether there was communication of the reconciled list back to patients, providers, and pharmacies.

Testing Objectives

As with all Alpha 2 data elements, the main testing objectives were to assess cross-setting feasibility and IRR (goals 1 and 2, as described in Chapter Two). The feedback that research nurses and facility/agency staff provided during the debriefing interview was used to gain an understanding of the feasibility of administering these items in PAC settings.

Results

Percentage of responses on the MR items, both overall and by setting, are presented in Table 8.1. Responses are also categorized by the type of assessor (i.e., research nurse or facility/agency staff). For descriptive purposes, we summarize results based on facility/staff ratings only in the remainder of this section.

Completion of item F1b was discontinued midway through the field period because of high time burden and because assessors confirmed with our team that the remainder of items would capture the MR process. For the subset of assessments that have a completed F1b, evidence for completion of MR was identified for one-half of the patients/residents overall, but this varied considerably by setting, with as few as 29 percent of IRF patients and as many as 75 percent of LTCH patients' records documenting evidence of a completed MR.

For F1c items, there was considerable variation in the overall percentage of patients/residents taking medications in each of the ten classes. Because of relatively infrequent responses for each drug class, the frequency of taking any medication within each class is reported (rather than the number of days taking a medication within each class). Nearly one-half of all patients/residents were receiving anticoagulants (47 percent) and opioids (45 percent); somewhat fewer patients/residents were receiving diuretics (37 percent), antidepressants (35 percent), antimicrobials (29 percent), and hypoglycemics (26 percent); less than 20 percent of patients/residents were receiving antianxiety medications (16 percent) and antiplatelets (15 percent); and only a handful of patients/residents were taking hypnotics (6 percent) and antipsychotics (4 percent). There was considerable variability in these percentages across settings as well. Indications for these drug classes tended to be recorded for about 50 percent of patients/residents, although indications were present for 50 of 53 patients/residents (94 percent) receiving opioids, 16 of 19 patients/residents (84 percent) receiving antianxiety medications, and 100 percent of patients/residents receiving hypnotics. As with the item that assessed what medications were being taken, documentation of the indications for these medications also varied

Table 8.1. Distribution of Responses to Medication Reconciliation Items, by Assessor and PAC Setting

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Documentation for complete MR (F1b) (percentage, *N*)	20, 2	53, 9	48, 10	29, 6	44, 4	75, 9	56, 10	55, 11	45, 26	50, 35
Anticoagulants (F1c1) (percentage, *N*)	33, 9	22, 6	58, 22	58, 22	79, 19	67, 16	44, 14	38, 11	53, 64	47, 55
d1: Indications (number)	4	5	7	11	6	3	7	7	24	26
e1: Discrepancies (number)	1	0	0	2	0	0	1	0	2	2
Antiplatelets (F1c2) (percentage, *N*)	7, 2	4, 1	29, 11	16, 6	21, 5	17, 4	19, 6	24, 7	20, 24	15, 18
d2: Indications (number)	1	1	8	2	1	0	4	6	14	9
e2: Discrepancies (number)	0	0	0	1	0	0	0	0	0	1
Hypoglycemics (F1c3) (percentage, *N*)	19, 5	11, 3	24, 9	21, 8	63, 15	54, 13	25, 8	24, 7	31, 37	26, 31
d3: Indications (number)	3	2	5	5	6	8	7	6	21	21
e3: Discrepancies (number)	0	0	1	4	0	0	1	0	2	4
Opioids (F1c4) (percentage, *N*)	33, 9	19, 5	68, 26	71, 27	50, 12	33, 8	72, 23	45, 13	58, 70	45, 53
d4: Indications (number)	6	4	25	27	11	8	21	11	63	50
e4: Discrepancies (number)	2	0	3	0	0	1	3	0	8	1
Antipsychotics (F1c5) (percentage, *N*)	7, 2	7, 2	5, 2	3, 1	8, 2	4, 1	3, 1	3, 1	6, 7	4, 5

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
d5: Indications (number)	1	1	0	0	2	0	1	0	4	2
e5: Discrepancies (number)	0	0	0	0	0	0	0	0	0	0
Antimicrobials (F1c6) (percentage, *N*)	19, 5	7, 2	18, 7	24, 9	67, 16	71, 17	25, 8	21, 6	30, 36	29, 34
d6: Indications (number)	2	2	3	6	4	6	7	5	16	19
e6: Discrepancies (number)	1	1	0	4	0	1	0	0	1	6
Antidepressants (F1c7) (percentage, *N*)	33, 9	33, 9	34, 13	24, 9	54, 13	50, 12	28, 9	38, 11	36, 44	35, 41
d7: Indications (number)	3	6	6	3	7	3	6	8	22	20
e7: Discrepancies (number)	1	0	0	3	0	0	0	0	1	3
Diuretics (F1c8) (percentage, *N*)	44, 12	33, 9	50, 19	24, 9	63, 15	46, 11	44, 14	52, 15	50, 60	37, 44
d8: Indications (number)	5	8	2	2	3	3	10	10	20	23
e8: Discrepancies (number)	3	0	2	3	0	0	1	1	6	4
Antianxiety (F1c9) (percentage, *N*)	11, 3	7, 2	26, 10	21, 8	25, 6	21, 5	19, 6	14, 4	21, 25	16, 19
d9: Indications (number)	1	2	8	8	5	2	5	4	19	16
e9: Discrepancies (number)	1	0	2	0	0	0	0	0	3	0
Hypnotics (F1c10) (percentage, *N*)	7, 2	4, 1	13, 5	11, 4	8, 2	0, 0	0, 0	7, 2	7, 9	6, 7
d10: Indications (number)	2	1	5	4	1	0	0	2	9	7
e10: Discrepancies (number)	0	0	0	0	0	0	0	0	0	0

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Was the reconciled medication list communicated to any of the following?[a] (F1i)										
F1i_1: Patient (percentage yes)	70	92	29	37	13	13	58	74	43	53
F1i_2: Prescribers/care providers (percentage yes)	59	72	100	97	100	83	97	78	90	84
F1i_3: Pharmacy (percentage yes)	15	12	100	76	91	39	94	56	77	50
F1i_4: None of the above (percentage yes)	15	4	0	0	0	13	0	11	3	6

NOTES: Data for F1c represent the percentage and number of patients/residents taking medications within a drug class in the past seven days or since admission. Data for F1d and F1e represent the number of patients/residents where the indication was noted for all medications in these medication classes and where there were discrepancies involving medications in these medication classes, respectively. These items are only relevant if the patient/resident took a medication in the medication class in the past seven days (F1c).

[a] Assessors may check all that apply, so these may add to more than 100 percent.

across settings. Overall, there were very few discrepancies identified for any of the drug classes; of those that were identified, the majority were appropriately addressed (data are not shown because of the small number of discrepancies identified). Finally, responses to the last item show that approximately one-half of patients/residents were provided with the final reconciled list. In contrast, a high percentage of assessments (84 percent) recorded that documentation showing the final list was communicated to the patient's/resident's prescriber and care team. Fifty percent of facility/agency staff assessments noted that the final list was communicated to the patient's/resident's primary pharmacy.

Feasibility

Table 8.2 shows the number of minutes taken to complete the MR section by research nurses and facility/agency staff, respectively. Across all settings combined, facility/agency staff took an average of 11.5 minutes, and research nurses took an average of 15.4 minutes to complete these items. As a reminder, we removed the F1b data element midway through the field period, meaning that times to complete in Table 8.2 are somewhat misleading: They are averaged across assessments completed with and without F1b. In general, the exclusion of F1b resulted in completion of the protocol about four and five minutes faster for facility/agency staff and research nurses, respectively.

Table 8.2. Mean Time Spent, in Minutes, Completing Medication Reconciliation Items, by PAC Setting

	HHA Mean (SD)	IRF Mean (SD)	LTCH Mean (SD)	SNF Mean (SD)	Overall Mean (SD)
By research nurse	9.4 (7.5)	15.8 (10.6)	20.1 (13.1)	16.7 (10.9)	15.4 (11.1)
By facility/agency staff	5.8 (4.3)	15.0 (9.2)	14.6 (6.6)	9.3 (10.6)	11.5 (9.0)

Interrater Reliability

Table 8.3 shows the IRR both overall and by setting. The IRR for F1b was calculated based on 56 research nurse and facility/agency staff pairs instead of 118 pairs because assessors were asked to stop collecting F1b in the middle of testing. Thus, 62 pairs have missing data because of deliberate skipping of F1b. The kappa for F1b was 0.32, indicating low agreement.

The IRRs for whether a patient was taking any medications within a class ranged from 0.33 to 0.88, but seven of the ten medication classes had an IRR of 0.65 or higher. IRRs for the indications and discrepancies data elements—which were only calculated overall, and not by setting because of low frequencies—were typically low but ranged considerably. The indications IRRs ranged from –0.50 to 0.73 and the IRRs for the discrepancies data elements ranged from –0.05 to 0.38. A negative kappa indicates that the two observers agreed less than would be expected just by chance.

The subsequent three items, F1f–F1h, on how the discrepancies were resolved, were skipped when there were no discrepancies identified in F1e. IRR was not calculated for these items because of sparse data in the 2 × 2 table used to calculate kappas. This was due, in part, to a low rate of discrepancies identified in F1e.

For communication of the reconciled medication list (F1i), the kappas were highest for communication with the patient (0.49) and lowest for communication to prescribers/care providers (–0.06).

Table 8.3. IRR of MR Items

	HHA	IRF	LTCH	SNF	Overall
Was complete MR done? (F1b) (yes/no)	0.29	0.22	–0.05	0.73	0.32
Anticoagulants (F1c1)	0.73	0.68	0.69	0.72	0.73
d1: Indications	—	—	—	—	0.40
e1: Discrepancies	—	—	—	—	N/A
Antiplatelets (F1c2)	–0.05	0.33	0.59	0.51	0.42
d2: Indications	—	—	—	—	0.07
e2: Discrepancies	—	—	—	—	N/A
Hypoglycemics (F1c3)	0.71	0.92	0.83	0.91	0.88
d3: Indications	—	—	—	—	0.25
e3: Discrepancies	—	—	—	—	0.26
Opioids (F1c4)	0.63	0.94	0.67	0.54	0.73
d4: Indications	—	—	—	—	–0.08
e4: Discrepancies	—	—	—	—	–0.04
Antipsychotics (F1c5)	0.46	0.65	0.65	1.00	0.65
d5: Indications	—	—	—	—	–0.50
e5: Discrepancies	—	—	—	—	N/A
Antimicrobials (F1c6)	0.52	0.68	0.9	0.44	0.72
d6: Indications	—	—	—	—	0.43
e6: Discrepancies	—	—	—	—	N/A
Antidepressants (F1c7)	1.00	0.75	0.75	0.85	0.83

	HHA	IRF	LTCH	SNF	Overall
d7: Indications	—	—	—	—	0.73
e7: Discrepancies	—	—	—	—	–0.05
Diuretics (F1c8)	0.77	0.47	0.67	0.86	0.68
d8: Indications	—	—	—	—	0.53
e8: Discrepancies	—	—	—	—	0.38
Antianxiety (F1c9)	0.78	0.42	0.65	0.52	0.55
d9: Indications	—	—	—	—	0.18
e9: Discrepancies	—	—	—	—	N/A
Hypnotics (F1c10)	–0.05	0.62	N/A	N/A	0.33
d10: Indications	—	—	—	—	N/A
e10: Discrepancies	—	—	—	—	N/A
Was the reconciled medication list communicated to any of the following? (F1i)					
F1i_1: Patient	–0.14	0.11	0.62	0.75	0.49
F1i_2: Prescribers/care providers	–0.32	N/A	N/A	–0.070	–0.06
F1i_3: Pharmacy	0.50	N/A	0.12	–0.15	0.33

NOTES: Sample size for F1b was 56 paired assessments because F1b was discontinued midway through the testing period because of high burden. For all other items, sample size was 118 paired assessments. Cells with N/A indicate that the frequency table is too sparse to compute IRR (i.e., not all response categories were endorsed by both nurses). Items F1f, F1g, F1h, and F1i_4 were not evaluated for IRR because of empty cells in the 2 × 2 tables used to construct kappas. Similarly, setting-specific kappas were not calculated for F1d and F1e because of insufficient data.

Feedback from Assessors

Although MR data elements were lengthy to complete and were cited as one of the most-complex domains of all those tested in Alpha 2, both research nurses and facility/agency staff noted that they became more efficient in completing the sections as they did more assessments. Facility/agency nurses may have taken somewhat less time to complete this section, on average, because of greater familiarity with the EHR. Research nurses noted that it took them time to become familiar with the provider's EHR and that interprovider variability in record management contributed to the challenge of completing the MR assessment. Assessors also noted that nurses may be better able to complete the MR assessment than other types of facility/agency staff because of their familiarity with medications.

Summary of Findings

Feasibility/ease of use: The time burden was reduced from 15 to 20 minutes in Alpha 1 (depending on facility/agency staff versus research nurses) to an average of eight to 12 minutes in Alpha 2 without F1b, despite adding much more detail in Alpha 2. Although the MR data element appears feasible to complete in all four settings, it continues to be a challenge for assessors. Feedback from assessors implies that the data-collection burden diminishes with experience.

Interrater reliability: Evidence for IRR of the MR data elements was mixed. For many items, IRR was moderate to high (0.65 or higher), but for others, IRR was unacceptably low. This was especially true for the indications and discrepancies items, which, with only two exceptions, had overall IRR values lower than 0.50. Because these items were only completed on the subset of patients/residents receiving each of the ten medication classes, the sample size for these items was small. Further testing on a larger sample is necessary to make strong conclusions about the performance of these items.

Recommendations

Substantial improvements have been made between the Alpha 1 and Alpha 2 testing. For example, the clarity of instructions appeared to improve from Alpha 1 to Alpha 2 because terms were defined within the item itself, rather than referring assessors to the user manual. In addition, many more coding examples were provided in the training sessions and user manual. However, results from Alpha 2 indicate several remaining limitations. Assessors found the item set overly burdensome and also reported confusion in collecting the data elements. Future development and testing efforts for these items are warranted to ensure good item performance.

Chapter Nine. Results for Noncommunicative Assessments

In this chapter, we focus on the three data elements developed for use with noncommunicative patients/residents. These elements are all collected via staff observation and include assessments of pain, mood, and cognitive status. The noncommunicative assessments were administered in the Alpha 2 test to a distinct sample of 44 patients/residents who met criteria, as previously described (see Table 3.1 in Chapter Three). This chapter is organized as follows:

- Observational Assessment of Pain
- Staff Assessment of Patient/Resident Mood
- Staff Assessment of Mental Status
- Overall Recommendations for Noncommunicative Assessments.

Each of the first three sections includes a description of the data elements, testing objectives, results, and a summary of findings. A set of recommendations that covers all three assessments is presented at the end of the chapter.

Observational Assessment of Pain

Pain affects a significant percentage of patients/residents in PAC settings.[1] Inattention to or mismanagement of pain can significantly affect care management and is associated with decreased quality of life, poor outcomes, and reduced participation in rehabilitation therapies.[2] Despite the fact that pain is a common and recognizable human experience, it is often underrecognized, underdetected, and understudied among older adults.[3] Evidence that rates of pain differ between groups based on cognitive status, age, and race implies that many PAC patients/residents who experience pain are not identified as such using current evaluation methods.[4]

---

[1] American Geriatrics Society Panel on the Pharmacological Management of Persistent Pain in Older Persons, "Pharmacological Management of Persistent Pain in Older Persons," *Pain Medication*, Vol. 10, No. 6, 2009, pp. 1062–1083.

[2] Nancy Wells, Chris Pasero, and Margo McCaffery, "Improving the Quality of Care Through Pain Assessment and Management," in R. G. Hughes, ed., *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*, Rockville, Md.: Agency for Healthcare Research and Quality, 2008.

[3] Alessio Avenanti, Domenica Bueti, Gaspare Galati, and Salvatore M. Aglioti, "Transcranial Magnetic Stimulation Highlights the Sensorimotor Side of Empathy for Pain," *Nature Neuroscience*, Vol. 8, No. 7, 2005, pp. 955–960.

[4] Stephen M. Thielke, Joanna Sale, and M. Carrington Reid, "Identifying, Tracking, and Managing Pain in LTC," *Annals of Long-Term Care*, Vol. 18, No. 9, 2010; Aza Abdulla, Nicola Adams, Margaret Bone, Alison M. Elliott, Jean Gaffin, Derek Jones, Roger Knaggs, Denis Martin, Liz Sampson, and Pat Schofield, "Guidance on the Management of Pain in Older People," *Age and Ageing*, Vol. 42, 2013, pp. i1–57.

Pain management can relieve symptoms, but accurate pain assessment is an essential precondition to managing pain. Assessment of pain helps to maintain standards of care and improve treatment planning and care management for patients/residents in PAC settings.[5] Because pain is a subjective experience for which there are no objective biological markers, self-report is often considered to be the gold standard for assessing pain. However, for noncommunicative patients/residents, interview assessment of pain is not possible; for these patients/residents, observational assessment of pain is a reasonable alternative.

Description of Data Element

The Observational Assessment of Pain data element can be found in Appendix B, Module E. Items collect the presence, over the past three days, of four behavioral pain indicators (nonverbal sounds, vocal complaints, facial expressions, body movements or postures, or none of these signs observed or documented). Then, for those exhibiting pain, the assessor was asked to record the frequency of the behaviors and whether indicators of pain resolved or diminished in response to pain medication or treatments.

Testing Objectives

As with all Alpha 2 data elements, the main testing objectives were to assess cross-setting feasibility and IRR (goals 1 and 2, as described in Chapter Two). Qualitative data obtained during the debriefing interviews were used to evaluate difficulties encountered in administering the observational pain assessment items.

Results

Table 9.1 shows the percentage of responses to each item, both overall and by setting. Responses are also categorized by the type of assessor (i.e., research nurse or facility/agency staff). For descriptive purposes, the remainder of this section summarizes results based on facility/agency staff ratings only.

There was considerable variability in the observation of pain indicators. For example, item E1a ("observed indicators of pain or distress") was coded as "none of these signs observed or documented" for both of the patients in the HHA setting, more than 50 percent of IRF patients, and exactly one-half of SNF residents. In contrast, item E1a was coded as "none of these signs observed or documented" for less than 30 percent of patients in the LTCH setting, suggesting that observed indicators of pain or distress are more prevalent among noncommunicative LTCH patients than patients/residents in other PAC settings. However, this variability in responses across settings was not surprising given the differences in patient/resident populations (e.g.,

[5] Wen-Chieh Lin, Terry Y. Lum, David R. Mehr, and Robert L. Kane, "Measuring Pain Presence and Intensity in Nursing Home Residents," *Journal of the American Medical Directors Association*, Vol. 7, No. 3, 2006, pp. 147–153.

presenting conditions, plans of care). Overall, slightly more than one-half (58 percent) of the noncommunicative patients/residents in the Alpha 2 sample exhibited one or more behaviors possibly indicative of pain. Among those exhibiting one or more behaviors, assessors observed a range in the frequency of the behaviors, with most occurring at least daily. The majority (74 percent) appeared to feel some relief with pain treatment, although this varied by setting, with 85 percent of LTCH patients and 75 percent of SNF residents but none of the five IRF patients evidencing some relief from pain treatment.

Table 9.1. Distributions of Responses to Observational Pain Items, by Assessors and PAC Setting

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Observed indicators of pain or distress (E1a)										
a. Nonverbal sounds (percentage)	0	0	10	11	29	25	33	20	24	20
b. Vocal complaints of pain (percentage)	0	0	10	11	18	8	33	40	18	16
c. Facial expressions (percentage)	0	0	30	33	50	33	22	20	37	29
d. Body movements or postures (percentage)	0	0	30	44	43	42	22	20	33	36
z. None of these signs observed/documented (percentage)	100	100	60	56	25	29	44	50	41	42
Missing (number)	0	0	0	0	0	0	0	0	0	0
Follow-up: Frequency of observed indicators and evidence that indicators diminished or were resolved with pain treatment *(percentages based on fewer than 30 patients/residents indicating pain or distress in E1a)*										
Frequency of observed indicators (E1b)										
Less than daily (percentage)	N/A	N/A	25	25	25	27	20	40	24	29
Daily (percentage)	N/A	N/A	25	25	30	20	60	40	34	25
More than daily (percentage)	N/A	N/A	50	50	45	53	20	20	41	46
Missing (number)	0	0	0	0	1	0	0	0	1	0
Unknown or unable to assess (number)	0	0	0	0	0	2	0	0	0	2
Evidence that indicators diminished/resolved with pain treatment (E1c)										
Yes (percentage)	N/A	N/A	33	0	93	85	100	75	86	74

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Missing (number)	0	0	1	1	4	1	2	1	7	3
Unknown or unable to assess (number)	0	0	0	1	2	3	0	0	2	4

NOTE: Cells with N/A indicate that the item was not completed, either because the assessors correctly skipped the item, indicated *unknown or unable to assess*, or responses were missing for all patients/residents.

Feasibility

Table 9.2 shows that the average time for facility/agency staff to complete the observational pain assessment was three minutes. Research nurses took more than a minute longer. The longer time taken by research nurses may be attributable to facility/agency staff having more-frequent opportunities for patient/resident observation as part of usual care during the assessment window, as well as greater familiarity with the patient's/resident's medical record at the facility/agency. There was some variation in completion time across settings. For example, assessment time was shorter in HHAs; however, this is likely attributable to the HHAs only having patients with no indicators of pain and therefore skipping two items. Overall, assessments of patients/residents with no indicators of pain took approximately 1.5 minutes less than those of patients/residents with indicators of pain. In addition, the skip rules were adhered to in all cases, but several responses were missing or "unknown or unable to assess" in item E1c (providing evidence that indicators diminished or resolved with pain treatment).

Table 9.2. Time Spent, in Minutes, Completing Observational Pain Items, by Assessor and PAC Setting

	HHA Mean (SD)	IRF Mean (SD)	LTCH Mean (SD)	SNF Mean (SD)	Overall Mean (SD)
By research nurse	1.75 (0.96)	4.50 (5.68)	5.19 (1.96)	4.67 (2.45)	4.68 (3.14)
By facility/agency staff	0.50 (0.71)	3.11 (2.32)	3.83 (2.60)	2.90 (2.47)	3.33 (2.51)

Interrater Reliability

Table 9.3 shows the IRR analysis of observational pain items. Overall kappas ranged from 0.69 to 1.00, indicating substantial to almost perfect interrater agreement for all items, which suggests that different assessors are able to obtain extremely similar results for these items.

Table 9.3. IRR of Observational Pain Items

	HHA	IRF	LTCH	SNF	Overall
Observed indicators of pain or distress[a] (E1a)					
a. Non-verbal sounds	N/A	1.00	0.60	0.73	0.69
b. Vocal complaints of pain	N/A	1.00	0.63	0.77	0.76
c. Facial expressions	N/A	1.00	0.67	1.00	0.80
d. Body movements or postures	N/A	0.77	0.83	1.00	0.85
z. None of these signs observed or documented	N/A	1.00	0.68	1.00	0.86
Frequency of observed indicators[b] (E1b)	N/A	1.00	0.92	0.74	0.90
Evidence that indicators diminished/resolved with pain treatment (E1c)[a]	N/A	N/A	1.00	N/A	1.00

NOTE: Cells with N/A indicate that the frequency table is too sparse to compute IRR (i.e., not all response categories were endorsed by both nurses).
[a] IRR was assessed by Cohen's kappa.
[b] IRR was assessed by weighted kappa.

Feedback from Assessors

Assessors noted that data collected from EHRs, staff, and direct observation of patients/residents were helpful in completing the observational pain items. The SNF and IRF assessors reported that it was easy to find documentation regarding pain because providers are focused on documenting pain, especially after hospitalization. Feedback from assessors suggested that talking to more than one staff member would be more helpful than relying on a report from only one.

One staff member mentioned that it was easy to check with staff about indicators of pain, but it was more difficult to coordinate a time to directly observe the patient/resident during an activity or while the patient is being turned.

Summary of Findings: Observational Assessment of Pain

Feasibility/ease of use: The pain items took approximately three minutes to complete; shorter completion times were reported among facility/agency staff than among research nurses. Assessors' comments generally reflected that the pain items were straightforward but somewhat challenging to administer because of the time required for observation and the need to consult multiple data sources.

Interrater reliability: Interrater agreement on observational pain items was substantial to almost perfect: 0.69 or above, with most items exceeding 0.80.

Staff Assessment of Patient/Resident Mood

In this section, we describe the observation-based assessment of mood items, testing objectives and analytic approach, and results from the feasibility test.

Description of Items

Items on the Staff Assessment of Patient/Resident Mood on the Patient Health Questionnaire (PHQ-9-OV) are presented in Appendix B, Module C. The PHQ-9-OV is designed to assess signs and symptoms of depressed mood in patients/residents who cannot complete a patient/resident mood interview because they are unable to communicate. The PHQ-9-OV assesses the nine signs and symptoms of depression included in the patient/resident interview–based PHQ-9, as well as irritability. Irritability is included because of its strong association with mood disorders in persons with cognitive impairment.

In completing the Staff Assessment of Patient/Resident Mood, assessors were instructed to interview staff members who know the patient/resident best, making an effort to interview staff from multiple shifts. In the case of home health care visits, assessors were asked to interview family members who have frequent contact with the patient. Assessors were also instructed to consult medical records covering the previous two weeks to look for clues about the patient's/resident's mood.

Each item on the Staff Assessment of Patient/Resident Mood has a symptom presence component and a symptom frequency component, the latter of which is administered only if a symptom is determined to be present. Frequency is quantified based on the number of days in the past 14 days the patient/resident has experienced this symptom. Possible frequency levels are "never or 1 day," "2–6 days (several days)," "7–11 days (half or more of the days)," and "12–14 days (nearly every day)." The assessors finalized the assessment by calculating a total score, which involves summing the symptom frequency ratings.

Testing Objectives

As with all Alpha 2 data elements, the main testing objectives were to assess cross-setting feasibility and IRR (goals 1 and 2, as described in Chapter Two). Qualitative data obtained during the debriefing interviews were used to evaluate difficulties encountered in administering the staff assessment of patient/resident mood items.

Results

Table 9.4 shows the frequency of responses for each item, both overall and by setting. Responses are also categorized by the type of assessor (i.e., research nurse or facility/agency staff member). For descriptive purposes, the remainder of this section summarizes results based on facility/agency staff ratings only. Between 40 percent and 50 percent of Alpha 2 patients/residents reported experiencing two cardinal symptoms of major depression, anhedonia

(C1a1), and depressed mood (C1b1), with anhedonia more apparent in SNFs and depressed mood more apparent in IRFs. Other symptoms of depressive disorders were also prevalent. More than 50 percent of patients/residents showed signs of being tired or having little energy, having trouble concentrating, and having psychomotor disturbances, and more than 40 percent showed signs of sleep disturbances and having a poor appetite or overeating.

Table 9.4. Distributions of Responses for the Staff Assessment of Patient/Resident Mood Items, by Assessors and PAC Setting

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Little interest or pleasure in doing things (C1a1)										
Yes (percentage)	N/A	N/A	44	38	31	41	71	70	44	49
Unknown or unable to assess (number)	4	2	1	1	11	7	2	0	18	10
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of “little interest or pleasure in doing things” (percentages based on fewer than 15 patients/residents indicating symptom presence in C1a1)										
If “yes” above, frequencies (C1a2)										
0. Never or 1 day (percentage)	N/A	N/A	0	0	0	0	0	20	0	8
1. 2–6 days (several days, percentage)	N/A	N/A	33	33	20	40	25	40	25	38
2. 7–11 days (half or more of the days, percentage)	N/A	N/A	33	33	40	40	0	0	25	23
3. 12–14 days (nearly every day, percentage)	N/A	N/A	33	33	40	20	75	40	50	31
Unknown or unable to assess (number)	0	0	1	0	0	2	1	2	2	4
Missing (number)	0	0	0	0	1	0	0	0	1	0
Feeling/appearing down, depressed, or hopeless (C1b1)										
Yes (percentage)	0	N/A	67	78	38	29	38	33	44	44
Unknown or unable to assess (number)	3	2	1	0	11	10	1	1	16	13
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of “feeling/appearing down, depressed or hopeless” (percentages based on fewer than 15 patients/residents indicating symptom presence in C1b1)										

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
If “yes” above, frequencies (C1b2)										
0. Never or 1 day (percentage)	N/A	N/A	17	14	0	0	0	0	8	8
1. 2–6 days (several days, percentage)	N/A	N/A	17	29	0	0	0	0	8	17
2. 7–11 days (half or more of the days, percentage)	N/A	N/A	33	14	20	33	50	50	31	25
3. 12–14 days (nearly every day, percentage)	N/A	N/A	33	43	80	67	50	50	54	50
Unknown or unable to assess (number)	0	0	0	0	1	1	1	1	2	2
Missing (number)	0	0	0	0	1	0	0	0	1	0
Sleep disturbances (C1c1)										
Yes (percentage)	N/A	N/A	30	33	52	59	57	56	48	53
Unknown or unable to assess (number)	4	2	0	0	2	2	2	1	8	5
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of sleep disturbances (percentages based on 20 or fewer patients/residents indicating symptom presence in C1c1)										
If “yes” above, frequencies (C1c2)										
0. Never or 1 day (percentage)	N/A	N/A	0	0	0	0	25	20	5	5
1. 2–6 days (several days, percentage)	N/A	N/A	33	33	38	17	0	20	30	20
2. 7–11 days (half or more of the days, percentage)	N/A	N/A	0	0	0	8	25	20	10	10
3. 12–14 days (nearly every day, percentage)	N/A	N/A	67	67	54	75	50	40	55	65
Unknown or unable to assess (number)	0	0	0	0	0	1	0	0	0	1
Missing (number)	0	0	0	0	1	0	0	0	1	0
Feeling tired or having little energy (C1d1)										
Yes (percentage)	100	N/A	50	44	79	79	71	78	71	70
Unknown or unable to assess (number)	2	2	0	0	8	5	2	1	12	8

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of feeling tired or having little energy (percentages based on 25 or fewer patients/residents indicating symptom presence in C1d1)										
If “yes” above, frequencies (C1d2)										
0. Never or 1 day (percentage)	0	N/A	20	25	7	8	0	17	8	13
1. 2–6 days (several days, percentage)	0	N/A	20	25	21	15	25	33	20	22
2. 7–11 days (half or more of the days, percentage)	0	N/A	40	25	0	0	0	0	8	4
3. 12–14 days (nearly every day, percentage)	100	N/A	20	25	71	77	75	50	64	61
Unknown or unable to assess (number)	0	0	0	0	1	2	1	1	2	3
Missing (number)	0	0	0	0	1	0	0	0	1	0
Poor appetite or overeating (C1e1)										
Yes (percentage)	25	N/A	67	43	40	38	71	44	53	42
Unknown or unable to assess (number)	0	2	1	2	17	16	2	1	20	21
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of poor appetite or overeating (percentages based on 15 or fewer patients/residents indicating symptom presence in C1e1)										
If “yes” above, frequencies (C1e2)										
0. Never or 1 day (percentage)	0	N/A	17	33	0	0	20	0	13	11
1. 2–6 days (several days, percentage)	0	N/A	0	0	0	0	20	25	7	11
2. 7–11 days (half or more of the days, percentage)	0	N/A	33	33	0	0	0	25	13	22
3. 12–14 days (nearly every day, percentage)	100	N/A	50	33	100	100	60	50	67	56
Unknown or unable to assess (number)	0	0	0	0	1	1	0	0	1	1
Missing (number)	0	0	0	0	1	0	0	0	1	0
Indicating s/he feels bad about self (C1f1)										
Yes (percentage)	N/A	N/A	40	33	13	14	38	33	24	24

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Unknown or unable to assess (number)	4	2	5	3	11	19	1	1	21	16
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of sleep disturbances (percentages based on fewer than 10 patients/residents indicating symptom presence in C1f1)										
If “yes” above, frequencies (C1f2)										
0. Never or 1 day (percentage)	N/A	N/A	0	0	50	50	33	33	29	29
1. 2–6 days (several days, percentage)	N/A	N/A	50	50	0	0	0	0	14	14
2. 7–11 days (half or more of the days, percentage)	N/A	N/A	0	0	0	0	33	33	14	14
3. 12–14 days (nearly every day, percentage)	N/A	N/A	50	50	50	50	33	33	43	43
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	1	0	0	0	1	0
Trouble concentrating (C1g1)										
Yes (percentage)	100	N/A	50	50	44	42	57	63	53	50
Unknown or unable to assess (number)	1	2	2	1	11	12	2	2	16	17
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of trouble concentrating (percentages based on fewer than 20 patients/residents indicating symptom presence in C1g1)										
If “yes” above, frequencies (C1g2)										
0. Never or 1 day (percentage)	0	N/A	0	0	0	0	0	0	0	0
1. 2–6 days (several days, percentage)	0	N/A	25	25	17	25	0	40	12	31
2. 7–11 days (half or more of the days, percentage)	0	N/A	0	0	17	0	0	20	6	8
3. 12–14 days (nearly every day, percentage)	100	N/A	75	75	67	75	100	40	82	62
Unknown or unable to assess (number)	0	0	0	0	1	1	0	0	1	1

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Missing (number)	0	0	0	0	1	0	0	0	1	0
Psychomotor disturbances (C1h1)										
Yes (percentage)	100	N/A	75	75	48	59	50	60	56	63
Unknown or unable to assess (number)	2	2	2	1	2	2	1	0	7	5
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of psychomotor disturbances (percentages based on 25 or fewer patients/residents indicating symptom presence in C1h1)										
If “yes” above, frequencies (C1h2)										
0. Never or 1 day (percentage)	0	N/A	20	0	0	0	0	0	4	0
1. 2–6 days (several days, percentage)	0	N/A	20	33	42	31	25	50	30	36
2. 7–11 days (half or more of the days, percentage)	0	N/A	0	0	17	31	25	33	13	24
3. 12–14 days (nearly every day, percentage)	100	N/A	60	67	42	38	50	17	52	40
Unknown or unable to assess (number)	0	0	1	0	0	0	0	0	1	0
Missing (number)	0	0	0	0	1	0	0	0	1	0
Thoughts of suicide or death (C1i1)										
Yes (percentage)	0	N/A	0	0	6	7	13	10	7	7
Unknown or unable to assess (number)	3	2	6	4	10	9	1	0	20	15
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of thoughts of suicide or death (percentages based on 2 or fewer patients/residents indicating symptom presence in C1i1)										
If “yes” above, frequencies (C1i2)										
0. Never or 1 day (percentage)	N/A	N/A	N/A	N/A	0	0	100	100	50	50
1. 2–6 days (several days, percentage)	N/A	N/A	N/A	N/A	0	0	0	0	0	0
2. 7–11 days (half or more of the days, percentage)	N/A	N/A	N/A	N/A	100	100	0	0	50	50

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
3. 12–14 days (nearly every day, percentage)	N/A	N/A	N/A	N/A	0	0	0	0	0	0
Unknown or unable to assess (number)	0	0	0	0	0	0	0	0	0	0
Missing (number)	0	0	0	0	1	0	0	0	1	0
Short-tempered or easily annoyed (C1j1)										
Yes (percentage)	50	100	40	44	32	26	25	30	34	33
Unknown or unable to assess (number)	0	1	0	0	5	5	1	0	6	6
Missing (number)	0	0	0	0	1	0	0	0	1	0
Follow-up: Frequency of being short-tempered or easily annoyed (percentages based on 15 or fewer patients/residents indicating symptom presence in C1j1)										
If "yes" above, frequencies (C1j2)										
0. Never or 1 day (percentage)	0	N/A	0	0	14	20	0	0	7	8
1. 2–6 days (several days, percentage)	0	N/A	0	25	14	0	50	67	13	25
2. 7–11 days (half or more of the days, percentage)	0	N/A	25	25	0	0	0	0	7	8
3. 12–14 days (nearly every day, percentage)	100	N/A	75	50	71	80	50	33	73	58
Unknown or unable to assess (number)	0	1	0	0	0	0	0	0	0	1
Missing (number)	0	0	0	0	1	0	0	0	1	0

NOTES: Responses may not sum to 100 percent because of rounding. Cells with N/A indicate that assessors correctly skipped the item, indicated *unknown or unable to assess*, or responses were missing for all patients/residents.

Feasibility

The rates of "unknown or unable to assess" in Table 9.4 were high for many of the depression symptoms, especially among LTCH and HHA patients. In contrast, assessors had little difficulty making assessments for IRF patients and SNF residents.

Table 9.5 shows that the average time to complete the Staff Assessment of Patient/Resident Mood was approximately 5.5 minutes for facility/agency staff. More than 70 percent of all assessments were completed in under nine minutes by facility/agency staff. Research nurses took

approximately two minutes longer to complete the items, and there was little variation across settings.

Table 9.5. Mean Time Spent Completing the Staff Assessment of Patient/Resident Mood

	HHA Mean (SD)	IRF Mean (SD)	LTCH Mean (SD)	SNF Mean (SD)	Overall Mean (SD)
By research nurse	5.75 (1.71)	7.30 (4.24)	7.59 (3.18)	7.67 (3.16)	7.40 (3.28)
By facility/agency staff	4.50 (0.71)	4.44 (3.81)	6.04 (3.11)	5.30 (3.27)	5.49 (3.22)

Interrater Reliability

As Table 9.6 shows, agreement between assessors in determining symptom presence was substantial for three items (feeling/appearing down, depressed, or hopeless, 0.72; sleep disturbances, 0.79; and poor appetite or overeating, 0.65) and nearly perfect for the other seven. IRR for symptom frequency was even better, with no value of Cohen's kappa lower than 0.81.

Table 9.6. IRR for the Items of the Staff Assessment of Patient/Resident Mood

	HHA	IRF	LTCH	SNF	Overall
Little interest or pleasure in doing things					
Presence (C1a1)	N/A	0.75	0.84	1.00	0.86
Frequency (C1a2)	—	—	—	—	N/A
Feeling/appearing down, depressed, or hopeless					
Presence (C1b1)	N/A	0.60	0.49	1.00	0.72
Frequency (C1b2)	—	—	—	—	0.90
Sleep disturbances					
Presence (C1c1)	N/A	1.00	0.72	0.70	0.79
Frequency (C1c2)	—	—	—	—	0.82
Feeling tired or having little energy					
Presence (C1d1)	N/A	0.78	1.00	0.59	0.86
Frequency (C1d2)	—	—	—	—	1.00
Poor appetite or overeating					
Presence (C1e1)	N/A	0.46	0.75	0.70	0.65
Frequency (C1e2)	—	—	—	—	0.89

	HHA	IRF	LTCH	SNF	Overall
Indicating patient/resident feels bad about self					
Presence (C1f1)	N/A	1.00	1.00	1.00	1.00
Frequency (C1f2)	—	—	—	—	1.00
Trouble concentrating					
Presence (C1g1)	N/A	0.72	1.00	0.67	0.83
Frequency (C1g2)	—	—	—	—	N/A
Psychomotor disturbances					
Presence (C1h1)	N/A	1.00	0.81	0.75	0.82
Frequency (C1h2)	—	—	—	—	N/A
Thoughts of suicide or death					
Presence (C1i1)	N/A	N/A	1.00	1.00	1.00
Frequency (C1i2)	—	—	—	—	1.00
Being short-tempered or easily annoyed					
Presence (C1j1)	N/A	1.00	0.87	1.00	0.94
Frequency (C1j2)	—	—	—	—	0.81

NOTES: IRR for presence items was assessed by Cohen's kappa. IRR for frequency items was assessed by weighted kappa. Cells with N/A indicate that the frequency table is too sparse to compute IRR (i.e., not all response categories were endorsed by both nurses).

Feedback from Assessors

Facility/agency staff and research nurses agreed that instructions for completing these items were straightforward. However, several concerns were mentioned, such as the 14-day look-back period. Facility/agency staff were concerned that the look-back period could be a challenge: Staff had a hard time answering the questions if the care of a patient/resident was only for one or two days. One also noted that a patient's mood can change frequently. Facility/agency staff reported that it is difficult to assess a patient's/resident's mood when the interaction is limited by the patient's/resident's noncommunicative status and when the patient/resident is bedridden. One assessor noted that it is particularly difficult for staff to assess whether a patient/resident "states that life isn't worth living, wishes for death, or attempts to harm self" among patients/residents who are not able to communicate. A research nurse also mentioned that the item assessing poor appetite or overeating was not applicable to this population in LTCHs because 75 percent of the patients were on tube feeding.

Summary of Findings: Staff Assessment of Patient/Resident Mood

Feasibility: The observational assessment of patient/resident mood does not seem to be more burdensome than the patient/resident interview-based mood assessment (PHQ-9, tested in Alpha 1), given that the times required to complete the assessments were similar (5.5 minutes by facility/agency staff for staff assessment of mood and 5.9 minutes for patient/resident interview). Given the percentages of "unknown or unable to assess" codes, it appears feasible to administer the observational assessment of mood in IRFs and SNFs (it is already part of the MDS 3.0) but perhaps not in LTCHs. Assessors noted that some items were difficult for staff to assess in LTCHs because the items were not applicable (e.g., poor appetite or overeating) or because of their inability to assess inner thoughts without any communication. The small number of HHAs included in this test makes it difficult to draw a conclusion about feasibility in that setting.

Interrater reliability: IRR was high for all items and varied little across settings.

Staff Assessment of Mental Status

In this final section, we describe the Staff Assessment of Mental Status data elements, the testing objectives, and results from the feasibility test.

Description of Items

The Staff Assessment of Mental Status is an observational assessment of long-term memory, short-term memory, memory/recall ability, and decisionmaking based on staff observation. As it is currently implemented in the MDS 3.0, staff complete this section of the assessment only if the resident cannot complete the BIMS. The item set is intended for use among patients/residents who are unable to communicate. Ratings of mental status are based on observation of the patient/resident, information provided by staff and family and friends, and medical records. The item set is currently used in the MDS 3.0. Item content can be found in Appendix B, section A1.

Testing Objectives

As with all Alpha 2 data elements, the main testing objectives were to assess cross-setting feasibility and IRR (goals 1 and 2, as described in Chapter Two). Qualitative data obtained during the debriefing interviews were used to evaluate difficulties encountered in administering the Staff Assessment of Mental Status items.

Results

Table 9.7 shows the frequencies for the Staff Assessment of Mental Status items, overall and broken down by setting. Responses are also categorized by the type of assessor (i.e., research nurse or facility/agency staff). For descriptive purposes, the remainder of this section summarizes results based on facility/agency staff ratings only. Overall, almost 70 percent of patients/residents were assessed as "short-term memory OK," and almost 50 percent were

assessed as "long-term memory OK." Rates of awareness of the current season and surroundings ranged from 21 percent to 60 percent. For cognitive skills for decisionmaking, 5 percent of patients/residents were assessed as independent, 5 percent as having modified independence, 30 percent as moderately impaired, and 60 percent as severely impaired.

Table 9.7. Distribution of Responses to Staff Assessment of Mental Status Items, by Assessors and PAC Setting

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Short-term memory (A1a)										
Memory OK (percentage)	100	100	67	80	57	50	89	89	70	67
Missing or unknown or unable to assess (percentage)	0	0	4	4	7	4	0	1	11	9
Long-term memory (A1b)										
Memory OK (percentage)	100	100	33	14	56	61	38	33	53	47
Missing or unknown or unable to assess (percentage)	0	0	4	2	10	6	1	1	15	9
Current season (A1ci)										
Yes (percentage)	0	N/A	80	50	25	28	33	30	32	31
Missing or unknown or unable to assess (percentage)	1	2	5	5	8	6	0	0	14	13
Location of own room (A1cii)										
Yes (percentage)	33	N/A	50	33	24	16	0	0	24	21
Missing or unknown or unable to assess (percentage)	1	2	4	3	7	5	1	2	13	12
Staff names and faces (A1ciii)										
Yes (percentage)	33	N/A	75	57	59	47	44	50	40	50
Missing or unknown or unable to assess (percentage)	1	2	2	2	6	5	0	2	9	11
Knows in facility/bed/home (A1civ)										
Yes (percentage)	33	N/A	86	67	64	70	33	33	59	60
Missing or unknown or unable to assess (percentage)	1	2	3	3	6	4	0	1	10	10
Cognitive skills for decisionmaking (A1d)										

	HHA		IRF		LTCH		SNF		Overall	
	R	F	R	F	R	F	R	F	R	F
Independent (percentage)	0	N/A	0	0	8	9	0	0	4	5
Modified independence (percentage)	0	N/A	10	0	4	5	12	10	6	5
Moderately impaired (percentage)	50	N/A	30	33	20	19	25	50	26	30
Severely impaired (percentage)	50	N/A	60	67	68	67	63	40	64	60
Missing or unknown or unable to assess (percentage)	0	2	0	0	3	3	1	0	4	5

NOTES: Responses may not sum to 100 percent because to rounding. Cells with N/A indicate that assessors correctly skipped the item, indicated *unknown or unable to assess*, or responses were missing for all patients/residents.

Feasibility

Table 9.8 shows the time, on average, to complete the Staff Assessment of Mental Status items. In general, this observational assessment tended to take longer to complete for research nurses (6.5 minutes) than facility/agency staff (3.9 minutes) and tended to take longer for patients/residents in LTCHs than in the other settings, perhaps because LTCH patients tend to be less mobile. Research nurses and facility/agency staff appeared to have difficulty assessing some of the items. They reported five to 15 patients/residents as "unknown or unable to assess" or missing across the items.

Table 9.8. Mean Time Spent, in Minutes, Completing the Staff Assessment of Mental Status Items by PAC Setting

	HHA Mean (SD)	IRF Mean (SD)	LTCH Mean (SD)	SNF Mean (SD)	Overall Mean (SD)
By research nurse	5.25 (1.50)	3.33 (1.80)	8.48 (5.48)	4.44 (2.83)	6.53 (4.83)
By facility/agency staff	1.00 (0.00)	3.78 (2.68)	4.83 (3.70)	2.50 (1.18)	3.91 (3.15)

Interrater Reliability

Table 9.9 shows interrater agreement. Interrater agreement ranged from 0.70 to 1.00 overall. Agreement tended to be higher in LTCH and SNF facilities and lowest for some items in IRFs (e.g., knowledge of staff names and faces). Sample size was small in HHAs, and, thus, IRR was not computed.

Table 9.9. IRR of the Staff Assessment of Mental Status Items

	HHA	IRF	LTCH	SNF	Overall
Short-term memory (A1a)	N/A	1.00	1.00	1.00	1.00
Long-term memory (A1b)	N/A	0.55	0.87	1.00	0.87
Current season (A1ci)	N/A	0.50	1.00	1.00	0.93
Location of own room (A1cii)	N/A	0.55	1.00	N/A	0.90
Staff names and faces (A1ciii)	N/A	0.09	0.79	1.00	0.70
Knows in facility/bed/home (A1civ)	N/A	0.57	1.00	1.00	0.94
Cognitive skills for decisionmaking (A1d)	N/A	1.00	1.00	0.83	0.96

NOTES: IRR for all items, except A1d, was assessed by Cohen's kappa. IRR for A1d was assessed by weighted kappa. Cells with N/A indicate that the frequency table is too sparse to compute IRR (i.e., not all response categories were endorsed by both nurses).

Feedback from Assessors

A research nurse noted that caregivers can be confused by *always/never* wording for the Staff Assessment of Mental Status. Despite instructions to consult with family members, caregivers, and staff, assessors said that in some cases they were unable to complete the items because of missing documentation.

Summary of Findings: The Staff Assessment of Mental Status

Feasibility: The assessment appears to be feasible to administer in all settings, although the average time to complete the assessment tended to be longer for patients/residents in LTCHs than in the other settings. In addition, there were several patients/residents with missing data or responses of "unknown or unable to assess," indicating that the assessors had difficulty determining answers to some questions for certain patients/residents.

Interrater reliability: For most settings, interrater agreement tended to be substantial. However, in IRF settings, reliability was lower, specifically for elements asking about knowledge of staff names and faces.

Overall Recommendations for Noncommunicative Assessments

Testing of the three noncommunicative assessments was completed for a relatively small number of patients/residents in this pilot test. At the time of this report, testing of these items will continue. However, the results from this test indicate that, overall, the IRR for all three assessments was quite good. Furthermore, results imply that, although all three assessments appeared to be feasible in SNFs and IRFs, they were somewhat more challenging and time-consuming to administer in LTCHs. The HHA sample was prohibitively small (N = 2 patients),

precluding any conclusions in that setting. Testing on a larger sample will also be needed to understand more about the administration challenges in LTCHs. The biggest difficulty relevant to all three assessments was that of deciding how to code certain items (like the presence of suicidal thoughts) without being able to communicate with the patient/resident. It is likely that this issue is somewhat exacerbated by the testing environment, and coding would be less challenging if it were conducted by facility/agency staff who were familiar with the patient/resident, along with other standardized data elements. Again, as we move through the development and testing phase of this work, we will continue to assess item limitations while filling the gaps in assessment.

Chapter Ten. Conclusion

In this chapter, we summarize the findings from Alpha 2 testing and provide recommendations for moving forward with data elements for standardized assessment in PAC settings.

Sample and Data Collection

A total of 14 PAC providers—three HHAs, four IRFs, three LTCHs, and four SNFs distributed across Denver, Houston, and Chicago—contributed assessments to the Alpha 2 test. Challenges were anticipated in reaching targeted assessment completion numbers, and efforts were made to recruit extra providers. Despite these efforts, while the number of paired communicative admission assessments conducted was only two fewer than the target number of 120, paired communicative assessments at discharge and noncommunicative assessments both fell well short of the goal.

Failure to meet target completion goals could be due to a variety of reasons, including not having enough English-speaking patients/residents (only English-speaking patients/residents were included in the testing), difficulty balancing the workflow of the assessments with regular care, and unforeseen absences from facility/agency staff assessors. Furthermore, research nurses and facility/agency staff noted several challenges that varied by type of assessment and PAC setting. For example, the general level of sickness in patients/residents, which tended to be higher in SNFs and LTCHs, may have played a role in the likelihood of successfully completing an assessment upon admission, as well as the length of stay and opportunities for assessment upon discharge. Some strategies to mitigate the risks for lower-than-expected assessment numbers were implemented midway through the Alpha 2 field period, and a complete list of recommended strategies, as documented in this report, will be considered for future testing.

Data Elements

Assessors collected data from communicative patients/residents in the categories of Cognitive Status, Behavioral Signs and Symptoms, Anxiety, Care Preferences, and MR. Some assessments were designed to be completed through patient/resident interview and others through review of the medical record. Data elements for use with noncommunicative patients/residents collected observational data on pain, mood, and mental status.

All data elements were evaluated by quantitative and qualitative means for feasibility of administration across PAC settings and IRR (i.e., consistency of ratings between two trained individuals, in this case a research nurse and a facility/agency staff member). Some data

elements were evaluated on additional criteria. In Table 10.1, we provide a summary of the results of this testing.

Table 10.1. Summary of Results for All Tested Data Elements

Data Element	Average Time to Complete (minutes)	Missing Data	Overall Reliability (IRR)	Comments
DOTPA items	5.0	Minimal	Mixed, 0.34–0.81	IRR was moderate overall but varied by item.
PASS	6.5	High	High, 0.78–0.92	IRR was more variable by setting (0.51–0.98).
Behavioral Signs and Symptoms	3.0	Minimal	N/A[a]	IRR was difficult to calculate because of the small number of behaviors exhibited; B1b, B1c, and B1k showed moderate to good reliability (0.77, 0.66, and 0.60, respectively).
PROMIS Anxiety	5.0	Minimal	Very high, 0.97–1.00	IRR was more variable by setting (0.80–1.00).
Physician Orders	4.0	Minimal	0.22–0.66	Time to complete is for Physician Orders and Goals of Care; information was difficult to locate; IRR was not calculated for many categories.
Goals of Care	4.0	Minimal	Poor, −0.35–0.49	Time to complete is for Physician Orders and Goals of Care; information was difficult to locate; IRR was not calculated for many categories.
MR	11.5	Minimal	Mixed, −0.50–0.88	For many items, IRR was moderate to high (0.65 or higher), but for others, IRR was unacceptably low.
Staff Assessment of Mental Status	4.0	Moderate	High, 0.70–1.00	IRR was more variable by setting (0.09–1.00).
Staff Assessment of Patient/Resident Mood	5.5	Minimal	High, 0.72–1.00	Responses were “unknown or unable to assess” for many patients/residents, especially among LTCHs and HHAs.
Observational Assessment of Pain	3.0	Minimal	High, 0.69–1.00	Several responses were missing or “unknown or unable to assess” in item E1c (evidence that indicators diminished/resolved with pain treatment).

[a] Reliability was not calculated for this data element.

Overall Feasibility

Data elements tested in Alpha 2 were generally feasible to administer. Feedback from assessors indicated that instructions for most items were clear and facilitated successful completion. Low levels of missing data for many of the items support this conclusion. While

data elements proved feasible to administer, the level of burden undertaken to complete the assessment varied across data elements and settings.

Time to complete data elements varied widely, with Behavioral Signs and Symptoms taking less than three minutes and Medication Reconciliation taking 12 minutes, on average, to complete. Length and complexity were cited as challenging factors in completing the MR section; however, feedback from assessors implied that the data-collection burden for this data element diminished with experience. While Care Preferences items were generally able to be completed, feedback from assessors indicated that it was difficult to locate the information needed to complete these items.

Some data elements proved to be more feasible in certain PAC settings than others. For example, the Staff Assessment of Patient/Resident Mood appears feasible to administer in IRFs and SNFs but perhaps not in LTCHs, while a larger sample of HHAs will be needed to draw conclusions about feasibility in that setting. For other data elements, the time to complete differed by setting. Assessment of cognitive status through both interview and observation (DOTPA, PASS, and Staff Assessment of Mental Status) tended to take longer to complete in LTCHs than in other settings. This was also the case for the assessment of Anxiety. The time to complete DOTPA and PASS, as well as logistical challenges and considerations of relevance, make them potentially problematic for administration in LTCHs and SNFs.

A challenge noted across several data elements (Behavioral Signs and Symptoms, Anxiety, and Staff Assessment of Patient/Resident Mood) was the look-back periods, which may be longer than the period of time that the assessor has cared for the patient/resident or the length of time that the patient/resident has been in the current setting. Look-back periods that extend prior to the date of admission may be difficult for the patient/resident to recall and may warrant reconsideration in future testing.

Overall IRR

Data elements tested in Alpha 2 generally exhibited good interrater agreement. Many data elements, including PASS, Anxiety, and the noncommunicative assessments, demonstrated substantial agreement. Reliability varied considerably across DOTPA items, with some items showing excellent reliability and others displaying unacceptably low reliability. Reliability was difficult to assess for some data elements, including several of the Behavioral items because of the very small number of behaviors that were exhibited. However, reliability was moderate to good for the three Behavioral items in which it could be assessed. Similarly, reliability was not able to be calculated for many Care Preferences items because of the infrequent nature of many categories of Physician Orders found in the medical records. For categories of Physician Orders where IRR could be calculated, it varied widely, both overall and within settings, while IRR for the Goals of Care items was generally poor. The reliability for MR was mixed. For many of these items, IRR was moderate to high, but for others, IRR was unacceptably low.

Recommendations

Based on the experience and feedback from the Alpha 2 field test, we have identified several recommendations that should be considered to maximize provider recruitment and assessment completion potential for future testing. First, we learned that is important to clarify information regarding provider recruitment and participant eligibility when reaching out to potential providers for participation. The patient/resident recruitment process may go more smoothly with project information that is more easily understood and more positively framed. The data-collection process will benefit from developing positive engagement with facility/agency staff and identifying how to capture discharges for assessment. Finally, additional examples of ways to introduce and explain the assessments to different types of patients to encourage participation will enhance field staff trainings for future tests.

In general, data elements were feasible to complete, although the length of time to administer some data elements may warrant enhancements to assessor training. IRR was good for most data elements, with a few exceptions. In some cases, testing on a larger sample will enable calculation of IRR for certain items and data elements that lacked data in this feasibility test. Future testing efforts would benefit from enhanced field staff training with additional example data-collection scenarios and role plays to increase understanding of and familiarity with the assessment among facility/agency staff assessors.

Appendix A. Current Item Use, Reliability/Validity, and Mode of Collection

Table A.1 presents the data elements used in Alpha 2 testing.

Table A.1. Data Elements in Alpha 2 Testing

Item Name	Current Assessment Instrument Use	Tested in PAC PRD	Reliability/Validity	Mode of Collection
Cognitive Function and Mental Status				
DOTPA CARE	N/A	No	The individual scales were tested as part of the AM-PAC assessment and showed high test-retest reliability (0.91–0.97), high subject-proxy reliability (0.68–0.90), high setting-specific intraclass correlation coefficients (0.82–0.93), and high internal-consistency reliability (Cronbach's alpha = 0.90–0.95).	Patient interview and observation
PASS	N/A	No	The PASS has been tested in older adult populations and has shown good discriminatory validity. Two studies of community-dwelling older adults found that patients with major cognitive impairment needed significantly more assistance ($F = 7.10$, $p = 0.009$) and had significantly lower adequacy scores ($p = 0.0095$) than individuals with normal cognition (American Geriatrics Society Panel on the Pharmacological Management of Persistent Pain in Older Persons, 2009; and Wells, Pasero, and McCaffery, 2008). The PASS Medication Management Task is also significantly correlated with the Global Cognitive Score ($r = -0.43$, $p < 0.0001$).	Patient interview
Behavioral Signs and Symptoms	MDS	Yes	The MDS 3.0 report noted that, in a sample of 349, assessment scores were extremely reliable, as indicated by the range of percentage agreement (0.912–1.000) and kappa = 0.90. Similarly, in a sample of 900 cases, reliability was high, as indicated by the range of percentage agreement (0.929–1.000) and kappa = 0.942.	Patient interview and observation
PROMIS Anxiety	N/A	No	PROMIS, an NIH Roadmap Initiative designed to improve self-reported outcomes, has developed and calibrated an item bank assessing anxiety. The 11 Anxiety items selected for Alpha 2 administration show high convergent validity with the general distress scale from the Mood and Anxiety Symptom Questionnaire ($r = 0.80$) and correlate highly ($r = 0.81$) with the depression item bank and the Center for Epidemiological Studies Depression scale ($r = 0.75$).	Patient interview

Item Name	Current Assessment Instrument Use	Tested in PAC PRD	Reliability/Validity	Mode of Collection
Noncommunicative				
Staff Assessment of Mental Status	MDS	Yes	Studies testing the Staff Assessment of Mental Status in nursing home patients have shown it to have good IRR (r = 0.80) and good validity based on its correlation with other assessments, such as the Blessed Test (r = 0.66, $p < 0.05$) and the Reisberg Global Deterioration Scale (r = 0.59, $p < 0.05$) (Casten et al., 1998; Lawton et al., 1998). The MDS report (Saliba and Buchanan, 2008) noted that in a sample of 349, reliability was high, as indicated by the range of percentage agreement (0.868–0.943) and kappa = 0.795. Similarly, in a sample of 900 cases, reliability was high, as indicted by the range of percentage agreement (0.896–0.983) and kappa = 0.900.	Observation
Staff Assessment of Patient/Resident Mood	MDS	Yes	The MDS report (Saliba and Buchanan, 2008) noted that in a sample of 349 SNF residents, assessment scores were extremely reliable, as indicated by the range of percentage agreement (0.96–1.000) and kappa = 0.873. Similarly, in a sample of 900 cases, reliability was high, as indicated by the range of percentage agreement (0.864–1.000) and kappa = 0.923.	Observation
Observational Assessment of Pain	N/A	No	N/A	Observation
Care Preferences				
Physician Orders	N/A	No	N/A	Chart review
Goals of Care	N/A	No	N/A	Patient interview

Item Name	Current Assessment Instrument Use	Tested in PAC PRD		Reliability/Validity	Mode of Collection
MR					
MR Protocol: days patient/resident received medications during last seven days/admission; indication noted for medications; medication discrepancies; discrepancies addressed by involving patient/family; discrepancies communicated to physician within 24 hours; recommended physician actions regarding discrepancies carried out; reconciled medication list communicated	N/A	No	N/A		Patient interview and chart review

NOTES: AM-PAC = Activity Measure for Post-Acute Care. PAC PRD = Post-Acute Care Payment Reform Demonstration.

Appendix B. Alpha 2 Assessment Protocol

MODULE A: COGNITION

A0. Brief Interview for Mental Status

[Patient/Resident]

A0a. Repetition of Three Words

ASK PATIENT/RESIDENT: "I am going to say three words for you to remember. Please repeat the words after I have said all three. The words are: sock, blue and bed. Now tell me the three words."

Number of words repeated by patient/resident after first attempt:

- ☐ 3 = Three
- ☐ 2 = Two
- ☐ 1 = One
- ☐ 0 = None or no answer

AFTER THE PATIENT'S/RESIDENT'S FIRST ATTEMPT, SAY: "I will repeat each of the three words with a cue and ask you about them later: sock, something to wear; blue, a color; bed, a piece of furniture." **You may repeat the words up to two more times.**

A0b. Year, Month, Day

ASK PATIENT/RESIDENT: "Please tell me what year it is right now."

Patient's/resident's answer is:

- ☐ 3 = Correct
- ☐ 2 = Missed by 1 year
- ☐ 1 = Missed by 2 to 5 years
- ☐ 0 = Missed by more than 5 years or no answer

A0c. **ASK PATIENT/RESIDENT:** "What month are we in right now?"

Patient's/resident's answer is:

- ☐ 2 = Accurate within 5 days
- ☐ 1 = Missed by 6 days to 1 month
- ☐ 0 = Missed by more than 1 month or no answer

A0d. **ASK PATIENT/RESIDENT:** "What day of the week is today?"

Patient's/resident's answer is:

- ☐ 1 = Accurate
- ☐ 0 = Incorrect or no answer

ASK PATIENT/RESIDENT: "Let's go back to the first question. What were those three words that I asked you to repeat?" **If unable to remember a word, give cue (i.e., something to wear; a color; a piece of furniture) for that word.**

A0e. Recalls "sock"?

- ☐ 2 = Yes, no cue required
- ☐ 1 = Yes, after cueing ("something to wear")
- ☐ 0 = No, could not recall or no answer

A0f. Recalls "blue"?

- ☐ 2 = Yes, no cue required
- ☐ 1 = Yes, after cueing ("a color")
- ☐ 0 = No, could not recall or no answer

A0g. Recalls "bed"?

- ☐ 2 = Yes, no cue required
- ☐ 1 = Yes, after cueing ("a piece of furniture")
- ☐ 0 = No, could not recall or no answer

A0-TIME

ASSESSOR: Enter your estimate of the minutes it took to complete this section.

A0-NOTES: Enter any notes for this section

A1. STAFF ASSESSMENT OF MENTAL STATUS

A1a. Short-Term Memory OK

Seems or appears to recall after 5 minutes

- ☐ 0 = Memory OK
- ☐ 1 = Memory problem
- ☐ 9 = Unknown or unable to assess

A1b. Long-Term Memory OK

Seems or appears to recall long past

- ☐ 0 = Memory OK
- ☐ 1 = Memory problem
- ☐ 9 = Unknown or unable to assess

A1c. Memory/Recall Ability: IS THE PATIENT/RESIDENT NORMALLY ABLE TO RECALL:

A1ci. Current season

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess

A1cii. Location of own room

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess

A1ciii. Staff names and faces

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess

A1civ. That he or she is in a nursing facility/hospital bed/rehabilitation facility/home

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess

A1d. Cognitive Skills for Daily Decisionmaking

Made decisions regarding tasks of daily life:

- ☐ 0 = Independent—decisions consistent/reasonable
- ☐ 1 = Modified independence—some difficulty in new situations only
- ☐ 2 = Moderately impaired—decisions poor; cues/supervision required
- ☐ 3 = Severely impaired—never/rarely made decisions
- ☐ 9 = Unknown or unable to assess

A1-TIME

ASSESSOR: Enter your estimate of the minutes it took to complete this section.

A1-NOTES: Enter any notes for this section

A4. PASS MEDICATION MANAGEMENT

[Patient/Resident]

SAY TO PATIENT/RESIDENT: The next task involves managing medications.

ASK PATIENT/RESIDENT: "Please read the prescription label and find the directions for taking this medication."

HAND PATIENT/RESIDENT FIRST BOTTLE OF MEDICATION AND WAIT UNTIL PATIENT/RESIDENT LOOKS UP

"If you were taking this medication today, when would you have to take the next pill?"

SUBTASK 1:

A4a. <u>Reports next time</u> first medication is to be taken <u>correctly</u> (based on testing time, matches direction on label)

No Assistance ☐

Verbal Assistance ☐ ☐ ☐

(Guiding or Directing Cues)

Visual Assistance ☐ ☐ ☐

(Gestures or Demonstration)

Physical Assistance ☐ ☐ ☐

(Tactile Cues, Physical Help)

88 = Not attempted ☐

(Due to environmental limitations or patient/resident safety)

ENTER SUBTASK 1; A4a SCORE ☐

ASK PATIENT/RESIDENT: "This medication organizer is like a pillbox. It has the days of the week across the top **[POINT]** and the time of the day **[POINT]** along the side. Using the organizer, distribute the pills to be taken tomorrow and the following day according to the directions on the prescription label **[PAUSE]**."

"Do you know what you are to do? Do you have everything that you need?"

WAIT FOR RESPONSE

SUBTASK 3:

A4c. Distributes pills from first pill bottle into correct time slots for the next 2 days (all pills & all slots indicated; days indicated)

No Assistance ☐

Verbal Assistance ☐ ☐ ☐

(Guiding or Directing Cues)

Visual Assistance ☐ ☐ ☐

(Gestures or Demonstration)

Physical Assistance ☐ ☐ ☐

(Tactile Cues, Physical Help)

88 = Not attempted ☐

(Due to environmental limitations or patient/resident safety)

ENTER SUBTASK 3; A4c SCORE ☐

ASK PATIENT/RESIDENT: "Now, please read the prescription label on this bottle and find the directions for taking this medication."

[HAND PATIENT/RESIDENT SECOND BOTTLE OF MEDICATION AND WAIT UNTIL PATIENT/RESIDENT LOOKS UP].

"If you were taking this medication today, when would you have to take the next pill?"

SUBTASK 4:

A4d. Reports next time second medication is to be taken correctly (based on testing time, matches direction on label)

No Assistance ☐

Verbal Assistance ☐ ☐ ☐

(Guiding or Directing Cues)

Visual Assistance ☐ ☐ ☐

(Gestures or Demonstration)

Physical Assistance ☐ ☐ ☐

(Tactile Cues, Physical Help)

88 = Not attempted ☐

(Due to environmental limitations or patient/resident safety)

ENTER SUBTASK 4; A4d SCORE ☐

ASK PATIENT/RESIDENT: "Again, using the organizer, distribute the pills to be taken tomorrow and the following day according to the prescription directions on the label. Do you know what you are to do?"

WAIT FOR RESPONSE

SUBTASK 6:

A4f. Distributes pills from second pill bottle into correct time slots for the next 2 days (all pills & all slots indicated; days indicated)

No Assistance ☐

Verbal Assistance ☐ ☐ ☐

(Guiding or Directing Cues)

Visual Assistance ☐ ☐ ☐

(Gestures or Demonstration)

Physical Assistance ☐ ☐ ☐

(Tactile Cues, Physical Help)

88 = Not attempted ☐

(Due to environmental limitations or patient/resident safety)

ENTER SUBTASK 6; A4f SCORE ☐

CALCULATE AND ENTER PASS MEDICATION MANAGEMENT INDEPENDENCE MEAN SCORE →

A4-TIME

ASSESSOR: Enter your estimate of the minutes it took to complete this section.

A4-NOTES: Enter any notes for this section

A5. DOTPA CARE-C

INSTRUCTIONS: All items in Section A5, DOTPA CARE-C, are based on staff/caregiver input or chart review. Do Not Ask Patient/Resident.

A5a. Does the patient/resident have any problems with memory, attention, problem-solving, planning, organizing, or judgment?

- ☐ 0 = No **[SKIP TO: A5-END TIME]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP TO: A5-END TIME]**

A5b. Please describe the patient's/resident's problems with the following: memory, attention, problem-solving, planning, organizing, and judgment.

- ☐ 0 = Mildly impaired: Demonstrates some difficulty with one or more of these cognitive abilities
- ☐ 1 = Moderately impaired: Demonstrates marked difficulty with one or more of these cognitive abilities
- ☐ 2 = Severely impaired: Demonstrates extreme difficulty with one or more of these cognitive abilities
- ☐ 9 = Unknown or unable to assess

A5c. How often is the patient/resident able to complete simple problems without assistance?

Simple problems: Following basic schedules; requesting assistance; using a call bell; identifying basic wants/needs; preparing a simple cold meal

Without Assistance: Patient performance without cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5d. How often is the patient/resident able to complete simple problems with assistance?

Simple problems: Following basic schedules; requesting assistance; using a call bell; identifying basic wants/needs; preparing a simple cold meal

With Assistance: Patient/resident performance with cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually

- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5e. How often is the patient/resident able to complete complex problems without assistance?

Complex problems: Working on a computer managing personal, medical, and financial affairs; preparing a complex hot meal; grocery shopping; route finding and map reading

Without Assistance: Patient/resident performance without cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5f. How often is the patient/resident able to complete complex problems with assistance?

Complex problems: Working on a computer managing personal, medical, and financial affairs; preparing a complex hot meal; grocery shopping; route finding and map reading

With Assistance: Patient/resident performance with cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5g. How often is the patient/resident able to recall basic information without assistance?

Basic Information: Personal information (e.g., family members, biographical information, physical location); basic schedules; names of familiar staff; location of therapy area

Without Assistance: Patient/resident performance without cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5h. How often is the patient/resident able to recall basic information with assistance?

Basic Information: Personal information (e.g., family members, biographical information, physical location); basic schedules; names of familiar staff; location of therapy area

With Assistance: Patient/resident performance with cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5i. How often is the patient/resident able to recall complex information without assistance?

Complex information: Complex and novel information (e.g., carry out multiple-step activities, follow a plan); anticipate future events (e.g., keeping appointments)

Without Assistance: Patient/resident performance without cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5j. How often is the patient/resident able to recall complex information with assistance?

Complex information: Complex and novel information (e.g., carry out multiple-step activities, follow a plan); anticipate future events (e.g., keeping appointments)

With Assistance: Patient/resident performance with cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5k. How often is the patient/resident able to complete simple activities without assistance?

Simple activities: Following simple directions; reading environmental signs or short newspaper/magazine/ book passage; eating a meal; completing personal hygiene; dressing

Without Assistance: Patient/resident performance without cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always

- ☐ 9 = Unknown or unable to assess

A5l. How often is the patient/resident able to complete simple activities with assistance?

Simple activities: Following simple directions; reading environmental signs or short newspaper/magazine/book passage; eating a meal; completing personal hygiene; dressing

With Assistance: Patient/resident performance with cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5m. How often is the patient/resident able to complete complex activities without assistance?

Complex activities: Watching a news program; reading a book; planning and preparing a meal; managing one's own medical, financial, and personal affairs

Without Assistance: Patient/resident performance without cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5n. How often is the patient/resident able to complete complex activities with assistance?

Complex activities: Watching a news program; reading a book; planning and preparing a meal; managing one's own medical, financial, and personal affairs

With Assistance: Patient/resident performance with cueing, assistive device, or other compensatory augmentative intervention

- ☐ 0 = Never or Rarely
- ☐ 1 = Sometimes
- ☐ 2 = Usually
- ☐ 3 = Always
- ☐ 9 = Unknown or unable to assess

A5-TIME

ASSESSOR: Enter your estimate of the minutes it took to complete this section.

A5-NOTES: Enter any notes for this section.

MODULE B: BEHAVIORAL SIGNS AND SYMPTOMS

B1. BEHAVIORAL SIGNS AND SYMPTOMS

All items in Module B: Behavioral Signs and Symptoms are based on staff/caregiver input or chart review. Do Not Ask Patient/Resident.

B1a. Physical behavioral symptoms directed toward others (e.g., hitting, kicking, pushing, scratching, grabbing, abusing others sexually)

- ☐ 0 = Behavior not exhibited
- ☐ 1 = Behavior of this type occurred 1 to 3 days
- ☐ 2 = Behavior of this type occurred 4 to 6 days, but less than daily
- ☐ 3 = Behavior of this type occurred daily
- ☐ 9 = Unknown or unable to assess

B1b. Verbal behavioral symptoms directed toward others (e.g., threatening others, screaming at others, cursing at others)

- ☐ 0 = Behavior not exhibited
- ☐ 1 = Behavior of this type occurred 1 to 3 days
- ☐ 2 = Behavior of this type occurred 4 to 6 days, but less than daily
- ☐ 3 = Behavior of this type occurred daily
- ☐ 9 = Unknown or unable to assess

B1c. Other behavioral symptoms not directed toward others (e.g., physical symptoms such as hitting or scratching self, pacing, rummaging, public sexual acts, disrobing in public, throwing or smearing food or bodily wastes, or verbal/vocal symptoms like screaming, disruptive sounds)

- ☐ 0 = Behavior not exhibited
- ☐ 1 = Behavior of this type occurred 1 to 3 days
- ☐ 2 = Behavior of this type occurred 4 to 6 days, but less than daily
- ☐ 3 = Behavior of this type occurred daily
- ☐ 9 = Unknown or unable to assess

If ALL responses to B1a, B1b, AND B1c are coded as either "behavior not exhibited" (0) or "unknown or unable to assess" (9), SKIP to B1K

IMPACT ON PATIENT/RESIDENT

Considering all the behavioral symptoms noted in B1a, B1b, and B1c, did any of the identified symptom(s):

B1e. Put the patient/resident at significant risk for physical illness or injury?

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess

B1f. Significantly interfere with the patient's/resident's care?

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess

B1g. Significantly interfere with the patient's/resident's participation in activities or social interaction?

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 8 = Not Applicable
- ☐ 9 = Unknown or unable to assess

B1h. Put others at significant risk for physical injury?

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess

B1i. Significantly intrude on the privacy or activity of others?

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess

B1j. Significantly disrupt the delivery of care or living environment of others?

- ☐ 0 = No
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess

REJECTION OF CARE

B1k. Did the patient/resident reject evaluation or care (e.g., bloodwork, taking medications, ADL assistance) that is offered by members of the care team or caregiver and necessary to achieve the patient's/resident's goals for health and well-being?

Do not include behaviors that have already been addressed (e.g., by discussion or care planning with the patient/resident or family), and determined to be consistent with patient/resident values, preferences, or goals.

- ☐ 0 = Behavior not exhibited

- ☐ 1 = Behavior of this type occurred 1 to 3 days
- ☐ 2 = Behavior of this type occurred 4 to 6 days, but less than daily
- ☐ 3 = Behavior of this type occurred daily
- ☐ 9 = Unknown or unable to assess

B-TIME

ASSESSOR: Enter your estimate of the minutes it took to complete this section.

B-NOTES: Enter any notes for this section.

MODULE C: MOOD

C1. STAFF ASSESSMENT OF PATIENT/RESIDENT MOOD (PHQ-9-OV©)

Over the last 2 weeks, did the patient/resident have any of the following problems or behaviors?

C1a1. SYMPTOM PRESENCE: Little interest or pleasure in doing things

- ☐ 0 = No **[SKIP TO C1B1]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP TO C1B1]**

C1a2. SYMPTOM FREQUENCY: Little interest or pleasure in doing things

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)
- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)
- ☐ 9 = Unknown or unable to assess

C1b1. SYMPTOM PRESENCE: Feeling or appearing down, depressed, or hopeless

- ☐ 0 = No **[SKIP TO C1C1]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP TO C1C1]**

C1b2. SYMPTOM FREQUENCY: Feeling or appearing down, depressed, or hopeless

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)
- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)
- ☐ 9 = Unknown or unable to assess

C1c1. SYMPTOM PRESENCE: Trouble falling or staying asleep, or sleeping too much

- ☐ 0 = No **[Skip to C1d1]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[Skip to C1d1]**

C1c2. SYMPTOM FREQUENCY: Trouble falling or staying asleep, or sleeping too much

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)
- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)

- ☐ 9 = Unknown or unable to assess

C1d1. SYMPTOM PRESENCE: Feeling tired or having little energy

- ☐ 0 = No **[SKIP to C1e1]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP to C1e1]**

C1d2. SYMPTOM FREQUENCY: Feeling tired or having little energy

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)
- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)
- ☐ 9 = Unknown or unable to assess

C1e1. SYMPTOM PRESENCE: Poor appetite or overeating

- ☐ 0 = No **[SKIP TO C1F1]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP TO C1F1]**

C1e2. SYMPTOM FREQUENCY: Poor appetite or overeating

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)
- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)
- ☐ 9 = Unknown or unable to assess

C1f1. SYMPTOM PRESENCE: Indicating that s/he feels bad about self, is a failure, or has let self or family down

- ☐ 0 = No **[SKIP TO C1G1]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP TO C1G1]**

C1f2. SYMPTOM FREQUENCY: Indicating that s/he feels bad about self, is a failure, or has let self or family down

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)
- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)
- ☐ 9 = Unknown or unable to assess

C1g1. SYMPTOM PRESENCE: Trouble concentrating on things, such as reading the newspaper or watching television

- ☐ 0 = No **[SKIP TO C1H1]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP TO C1H1]**

C1g2. SYMPTOM FREQUENCY: Trouble concentrating on things, such as reading the newspaper or watching television

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)
- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)
- ☐ 9 = Unknown or unable to assess

C1h1. SYMPTOM PRESENCE: Moving or speaking so slowly that other people have noticed. Or the opposite, being so fidgety or restless that s/he has been moving around a lot more than usual

- ☐ 0 = No **[SKIP TO C1I1]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP TO C1I1]**

C1h2. SYMPTOM FREQUENCY: Moving or speaking so slowly that other people have noticed. Or the opposite, being so fidgety or restless that s/he has been moving around a lot more than usual

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)
- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)
- ☐ 9 = Unknown or unable to assess

C1i1. SYMPTOM PRESENCE: States that life isn't worth living, wishes for death, or attempts to harm self

- ☐ 0 = No **[SKIP TO C1J1]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP TO C1J1]**

C1i2. SYMPTOM FREQUENCY: States that life isn't worth living, wishes for death, or attempts to harm self

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)

- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)
- ☐ 9 = Unknown or unable to assess

C1j1. SYMPTOM PRESENCE: Being short-tempered, easily annoyed

- ☐ 0 = No **[SKIP TO TOTAL SCORE]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[SKIP TO TOTAL SCORE]**

C1j2. SYMPTOM FREQUENCY: Being short-tempered, easily annoyed

- ☐ 0 = Never or 1 day
- ☐ 1 = 2–6 days (several days)
- ☐ 2 = 7–11 days (half or more of the days)
- ☐ 3 = 12–14 days (nearly every day)
- ☐ 9 = Unknown or unable to assess

PHQ-9-OV TOTAL: Add values from C1a2, C1b2, C1c2, C1d2, C1e2, C1f2, C1g2, C1h2, C1i2 and C1j2 →

C-TIME

ASSESSOR: Enter your estimate of the minutes it took to complete this section.

C-NOTES: Enter any notes for this section.

MODULE D. ANXIETY

[Patient/Resident]

D1. PROMIS®—ANXIETY

[Patient/ Resident]

SAY TO PATIENT/RESIDENT: "I am now going to ask you about your emotional distress, specifically anxiety and how you have been feeling over the past 7 days. I will also ask about some common problems that sometimes go along with feeling anxious. This is not meant to give you a diagnosis. Some of the questions might seem personal, but all patients/residents are asked to answer them. Knowing the answers to these questions will help us provide you with a more individualized care plan."

D1a. In the past 7 days, I had difficulty sleeping

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1b. In the past 7 days, I felt worried

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1c. In the past 7 days, my worries overwhelmed me

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1d. In the past 7 days, I had trouble paying attention

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1e. In the past 7 days, I felt nervous

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1f. In the past 7 days, I felt anxious

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1g. In the past 7 days, I had difficulty calming down

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1h. In the past 7 days, I had a racing or pounding heart

- ☐ 1 = Never
- ☐ 2 = Rarely

- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1i. In the past 7 days, I found it hard to focus on anything other than my anxiety

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1j. In the past 7 days, I felt like I needed help for my anxiety

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D1k. In the past 7 days, I had sudden feelings of panic

- ☐ 1 = Never
- ☐ 2 = Rarely
- ☐ 3 = Sometimes
- ☐ 4 = Often
- ☐ 5 = Always
- ☐ 7 = Patient/resident declined to respond
- ☐ 9 = Unknown or unable to assess

D-TIME

ASSESSOR: Enter your estimate of the minutes it took to complete this section.

D-NOTES: Enter any notes for this section

MODULE E. PAIN

E1. PAIN

E1a. OBSERVATIONAL ASSESSMENT OF PAIN OR DISTRESS

For all patients/residents who are unable to participate in the pain interview, please note whether any of the following behaviors were observed.

Patients/residents should be observed twice daily (morning AND evening) during care activities (i.e., during transfer procedures, repositioning, bathing, toileting, wound care/dressing changes, range of motion, ambulating, or other exercises), when behavioral signs of potential pain or distress are most likely to be expressed, over the course of 3 consecutive days.

CHECK ALL THAT APPLY

- ☐ a = Nonverbal sounds (e.g., crying, whining, gasping, moaning, or groaning)
- ☐ b = Vocal complaints of pain (e.g., "that hurts, ouch, stop")
- ☐ c = Facial expressions (e.g., grimaces, winces, wrinkled forehead, furrowed brow, clenched teeth or jaw, rapid eye blinking, tightly closed eyes)
- ☐ d = Body movements or postures (e.g., bracing, guarding, rubbing or massaging a body part/area, clutching or holding a body part during movement, rigid, tense body posture; withdrawing an extremity to an external stimulus; fidgeting; increased pacing, rocking; restricted movement; gait or mobility changes)
- ☐ z = None of these signs observed or documented. **[SKIP TO E-END TIME]**

E1b. For patients/residents who demonstrated any indicators of potential pain or distress listed in E1a (*Observational Assessment of Pain or Distress*), identify the frequency with which patient complains or shows evidence of potential pain or distress over the past 3 days.

- ☐ 1 = Indicators of potential pain or distress observed less than daily
- ☐ 2 = Indicators of potential pain or distress observed daily (at least once per day on each day of the assessment window)
- ☐ 3 = Indicators of potential pain or distress observed more than daily (multiple times per day on each day of the assessment window)
- ☐ 9 = Unknown or unable to assess

E1c. For patients/residents who demonstrated any indicators of potential pain or distress listed in E1a (*Observational Assessment of Pain or Distress*), is there any evidence that these indicators resolved or diminished in response to pain medications or treatments over the past 3 days?

- ☐ 0 = No
- ☐ 1 = Yes

- ☐ 8 = Not applicable—patient/resident has not received pain medications or treatments within the past 3 days
- ☐ 9 = Unknown or unable to assess

E-TIME

ASSESSOR: Enter your estimate of minutes it took to complete this section.

E-NOTES: Enter any notes for this section.

MODULE F. MEDICATION RECONCILIATION

F1. Medication Reconciliation

All items in Module F: Medication Reconciliation are based on staff/caregiver input, chart review, or communication from the patient/resident.

F1b. Is there documentation that complete medication reconciliation was done?

- ☐ 0 = No
- ☐ 1 = Yes

F1c. Indicate the number of DAYS the patient/resident received the following medications during the last seven days or since admission/discharge/SOC/ROC if less than seven days. If the patient/resident is taking more than one medication in the same class, the highest number of days should be used.

F1C1: Anticoagulants	DAYS:
F1C2: Antiplatelets (excluding 81 mg aspirin)	DAYS:
F1C3: Hypoglycemics (for example, insulin)	DAYS:
F1C4: Opioids	DAYS:
F1C5: Antipsychotics	DAYS:
F1C6: Antimicrobials (excluding topicals)	DAYS:
F1C7: Antidepressants	DAYS:
F1C8: Diuretics	DAYS:
F1C9: Antianxiety	DAYS:
F1C10: Hypnotics	DAYS:

F1d. Was there an indication noted for all medications in these medication classes?

CHECK ONE BOX FOR EACH OF THE MEDICATION CLASSES THE PATIENT/RESIDENT IS TAKING

	NO (0)	YES (1)
F1d1: Anticoagulants	☐	☐
F1d2: Antiplatelets (excluding 81 mg aspirin)	☐	☐
F1d3: Hypoglycemics (for example, insulin)	☐	☐
F1d4: Opioids	☐	☐
F1d5: Antipsychotics	☐	☐
F1d6: Antimicrobials (excluding topicals)	☐	☐
F1d7: Antidepressants	☐	☐
F1d8: Diuretics	☐	☐
F1d9: Antianxiety	☐	☐
F1d10: Hypnotics	☐	☐

F1e. Were there discrepancies involving medications in these medication classes?

CHECK ONE BOX FOR EACH OF THE MEDICATION CLASSES THE PATIENT/RESIDENT IS TAKING

	NO (0)	YES (1)	Missing information on sources OR lack of documentation (9)
F1e1: Anticoagulants	☐	☐	☐
F1e2: Antiplatelets (excluding 81 mg aspirin)	☐	☐	☐
F1e3: Hypoglycemics (for example, insulin)	☐	☐	☐
F1e4: Opioids	☐	☐	☐
F1e5: Antipsychotics	☐	☐	☐
F1e6: Antimicrobials (excluding topicals)	☐	☐	☐
F1e7: Antidepressants	☐	☐	☐
F1e8: Diuretics	☐	☐	☐
F1e9: Antianxiety	☐	☐	☐
F1e10: Hypnotics	☐	☐	☐

F1f. Were the patient's/resident's discrepancies regarding these medication classes addressed by involving the patient/resident or patient's/resident's family/formal caregiver?

CHECK ONE BOX FOR EACH OF THE MEDICATION CLASSES THE PATIENT/RESIDENT IS TAKING

	NO (0)	YES (1)	Missing information on sources OR lack of documentation (9)
F1f1: Anticoagulants	☐	☐	☐
F1f2: Antiplatelets (excluding 81 mg aspirin)	☐	☐	☐
F1f3: Hypoglycemics (for example, insulin)	☐	☐	☐
F1f4: Opioids	☐	☐	☐
F1f5: Antipsychotics	☐	☐	☐
F1f6: Antimicrobials (excluding topicals)	☐	☐	☐
F1f7: Antidepressants	☐	☐	☐
F1f8: Diuretics	☐	☐	☐
F1f9: Antianxiety	☐	☐	☐
F1f10: Hypnotics	☐	☐	☐

F1g. Were discrepancies regarding these medication classes communicated to the physician (or physician-designee) within 24 hours of admission/discharge/SOC/ROC?

CHECK ONE BOX FOR EACH OF THE MEDICATION CLASSES THE PATIENT/RESIDENT IS TAKING

	NO (0)	YES (1)	Missing information on sources OR lack of documentation (9)
F1g1: Anticoagulants	☐	☐	☐
F1g2: Antiplatelets (excluding 81 mg aspirin)	☐	☐	☐
F1g3: Hypoglycemics (for example, insulin)	☐	☐	☐
F1g4: Opioids	☐	☐	☐
F1g5: Antipsychotics	☐	☐	☐
F1g6: Antimicrobials (excluding topicals)	☐	☐	☐
F1g7: Antidepressants	☐	☐	☐
F1g8: Diuretics	☐	☐	☐
F1g9: Antianxiety	☐	☐	☐
F1g10: Hypnotics	☐	☐	☐

F1h. Were recommended physician (or physician-designee) actions regarding discrepancies for these medication classes carried out within 24 hours after the physician responded?

CHECK ONE BOX FOR EACH OF THE MEDICATION CLASSES THE PATIENT/RESIDENT IS TAKING

	NO (0)	YES (1)	Physician has not responded (8)	Missing information on sources OR lack of documentation (9)
F1h1: Anticoagulants	☐	☐	☐	☐
F1h2: Antiplatelets (excluding 81 mg aspirin)	☐	☐	☐	☐
F1h3: Hypoglycemics (for example, insulin)	☐	☐	☐	☐
F1h4: Opioids	☐	☐	☐	☐
F1h5: Antipsychotics	☐	☐	☐	☐
F1h6: Antimicrobials (excluding topicals)	☐	☐	☐	☐
F1h7: Antidepressants	☐	☐	☐	☐
F1h8: Diuretics	☐	☐	☐	☐
F1h9: Antianxiety	☐	☐	☐	☐
F1h10: Hypnotics	☐	☐	☐	☐

F1i. Was the reconciled medication list communicated to any of the following?

CHECK ALL THAT APPLY

- ☐ 1 = Patient/resident or patient's/resident's family/formal caregiver
- ☐ 2 = Prescribers and the care team responsible for the patient's/resident's care following admission/discharge/SOC/ROC
- ☐ 3 = Patient's/resident's pharmacy that will be filling most of the medications following admission/discharge/SOC/ROC
- ☐ 9 = Missing information sources or lack of documentation

F-TIME

ASSESSOR: Enter your estimate of the minutes it took to complete this section.

F-NOTES: Enter any notes for this section.

MODULE G. CARE PREFERENCES

G1. Advance Directives

All items in Section G1 (G1a–g) are based on chart review only. Do Not Ask Patient/Resident.

G1b. Does the patient/resident have any of the following physician orders documented and active in the medical record?

CHECK ALL THAT APPLY:

- ☐ a = Do not resuscitate (DNR)
- ☐ b = Do not intubate (DNI)
- ☐ c = Do not hospitalize (DNH)
- ☐ d = Antibiotic restrictions
- ☐ e = Comfort care preference(s)
- ☐ z = None of the above

G1c. Is there documentation in the medical record indicating that a conversation between the patient/resident (or representative) and the care team (or physician) took place about the patient's/resident's goals for care?

- ☐ 0 = No **[IF NO, SKIP TO G1-END TIME]**
- ☐ 1 = Yes
- ☐ 9 = Unknown or unable to assess **[IF NO, SKIP TO G1-END TIME]**

G1d. Did the documented conversation about goals of care indicate any of the following types of goals?

CHECK ALL THAT APPLY:

- ☐ 1 = Physical Goals
- ☐ 2 = Emotional Goals
- ☐ 3 = Social Goals
- ☐ 4 = Intellectual/Mental Goals
- ☐ 5 = Other: _____

G1-TIME

ASSESSOR: Enter your estimate of the minutes it took to complete this section.

G1-NOTES: Enter any notes for this section

H. FINAL CHECKS

[Assessor Only]

H1. ***ASSESSOR:*** *Is this an incomplete assessment due to patient/resident transfer or death?*

- ☐ 0 = No
- ☐ 1 = Yes

H1-NOTE:

H2. ***ASSESSOR:*** *Did the patient/resident opt out of the study during the assessment?*

- ☐ 0 = No
- ☐ 1 = Yes

H2-NOTE:

H3. ***ASSESSOR:*** *Describe any problems experienced using the tablet to conduct the assessment:*

INSTRUCTIONS TO ASSESSOR:

Please check the crosswalk and all assessment modules to verify this assessment is complete and ready to be finalized.

Appendix C. Assessor Debrief Interview Protocol

Interview Protocol: Research Nurses

Alpha 2 Test Debrief Interview for Research Nurses

Date:
Attendees:
General notes:
Action items:

Assessment Progress

Question: To begin, to date how many assessments have each of you conducted and in what settings?

Assessment Overview

Question: Thinking of these assessments, can someone describe an assessment that went really smoothly? What do you think made this assessment easy to conduct/complete?

Question: Now I want you to think about a fairly challenging assessment you've completed with your partner. What do you think made this assessment so challenging to conduct/complete?

Follow up: What steps did you and your partner have to take to address these challenges? Or, what steps should be taken in the future to address these types of challenges?

Follow up: Were any of these challenges unique to the setting (i.e., IRF, LTCH, HHA, SNF)?

User Manual

Question: When did you find yourself referencing the user manual?

Follow up: To what extent was [the user manual] helpful? Confusing? Do you have any suggestions for improvements or additions regarding the user manual?

Follow up: Do you refer to the user manual more for observation-based assessments versus interview-based assessments?

Tablet

Question: To what extent have you encountered challenges with the tablet's performance?

Question: How easy is it to use the tablet? What is frustrating about using it?

Question: What improvements do you suggest?

Question: What are the best aspects of using the tablet?

Question: How is it different than using paper?

Question: What are your thoughts about the crosswalk interface? How is the workflow of switching back and forth to get the ID number? Any other thoughts about the crosswalk? How can it be improved?

Training follow up: Did you feel like the training you previously received on using the tablet was helpful? What could have been done to improve the training?

Do you feel the facility staff was well prepared to use the tablets to collect data during the assessments? What challenges did they face?

Training follow up: What could have been done to improve the training that facility staff received?

Facility Staff

Question: How comfortable have they been when conducting assessments? What challenges have they faced when conducting the assessments?

Training follow up: What, if any, additional tools/materials would have helped you train the facility staff during the training that you attended in Chicago, Denver, and Houston?

Training follow up: Looking back at the field training you provided the field staff, given what you have experienced in conducting the assessments so far, what would you recommend we change about field training?

Cognition

General comments

Question: Please describe any challenges you encountered that you think were setting-specific.

Question: Describe any challenges in obtaining the information to complete this domain.

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Behavioral Signs and Symptoms

General comments

Question: What helped with completing these data elements (i.e., which information sources/approaches worked best)?

Question: Describe any challenges you encountered that seemed to be setting-specific.

Question: How were the challenges overcome?

Question: Describe any challenges in obtaining the information to complete this domain. How were the challenges overcome?

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Question: Do you have any suggestions for improving the training or clarity of instructions for this section?

Mood

General comments

Straightforward

Hard-to-answer questions that were about themselves. Feel depressed, etc.

Question: How often were you uncertain about whether or how often a patient/resident exhibited one of the symptoms of depression that are included in the assessment? Were there particular symptoms that you had more trouble making decisions about than other symptoms? If so, which ones gave you the most trouble?

Question: What sources of information did you rely on for completing this assessment?

Question: Describe any challenges that you had in obtaining the information needed to complete this assessment.

Question: Describe any challenges that you encountered that seemed to be specific to a particular setting.

Question: How easy or difficult do you think it will be to gather the information needed to complete this assessment in practice (i.e., outside of a testing situation)? What might some of the barriers be when completing this assessment in practice?

Anxiety

General comments

Question: Describe any challenges you encountered that seemed to be setting-specific.

Question: Describe any challenges in obtaining the information to complete this domain.

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Question: Did the patients express any concern about their privacy when asked their anxiety symptoms?

Question: After you inform the patients that these questions are not meant to give them a diagnosis, but to assist planning for better care, did patients feel comfortable talking about their anxiety?

Question: Did patients have any trouble recalling their symptoms during a lookback period of 7 days?

Pain

General comments

Question: Describe any challenges you encountered that seemed to be setting-specific.

Question: Describe any challenges in obtaining the information to complete this domain.

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Medication Reconciliation

General comments

Question: Describe any challenges you encountered that seemed to be setting-specific.

Question: Describe any challenges in obtaining the information to complete this domain.

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Med Rec–Specific Question: Was enough information available to answer? Do you think people were clear on when to answer no versus 9?

Med Rec Tablet Question: Did the format work? Built-in logic did not allow all rows to show up—how what that?

Care Preferences

General comments

Describe any challenges you encountered that seemed to be setting-specific.

Describe any challenges in obtaining the information to complete this domain.

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Follow up question: Did you understand what we meant by "Goals of Care" and the different areas of "Goals of Care" (i.e., physical goals, social goals, intellectual goals, emotional goals)?

Question: To what extent did the EHR help/hinder you in completing this section?

Sources of Information

In general, for domains or items that require multiple sources of information to complete the assessment, what sources did you tend to use or have access to? What types of challenges did you encounter with these questions?

General Comments on Entire Process

Do you have any other observations/thoughts about the process that you would like to share?

Action Items:

Note Taker Comments:

Changes to Interview Structure:

Interview Protocol: Facility/Agency Staff

Alpha 2 Test Debrief Interview for Field Staff

Date:
Attendees:
General notes:
Action items:

Assessment Progress

To begin, please describe an assessment that you have done since starting the pilot [pilot is only used for the first focus group conducted between Weeks 2 and 3] data collection that you think went smoothly. In what ways did it run well; what made it easier than others to conduct/complete?

Assessment Overview

Next, please describe an assessment that was particularly challenging to conduct/complete. What do you think made it so challenging? What steps were taken, or should be taken, to mitigate these challenges?

Thinking across all the pilot [pilot is only used for the first focus group conducted between Weeks 2 and 3] assessments you have completed thus far, tell us some more about what challenges you have encountered: Probes: Patient characteristics, time, data entry process, coding/scoring, the assessment tool itself, physical space and environmental issues (e.g., lack of privacy, interruptions, background noise).

Cognition

General comments

Question: Please describe any challenges you encountered that you think were setting-specific.

Question: Describe any challenges in obtaining the information to complete this domain.

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Behavioral Signs and Symptoms

General comments

Question: What helped with completing these data elements (i.e., which information sources/approaches worked best)?

Question: Describe any challenges you encountered that seemed to be setting-specific.

Question: How were the challenges overcome?

Question: Describe any challenges in obtaining the information to complete this domain. How were the challenges overcome?

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Question: Do you have any suggestions for improving the training or clarity of instructions for this section?

Mood

General comments

Question: How often were you uncertain about whether or how often a patient/resident exhibited one of the symptoms of depression that are included in the assessment? Were there

particular symptoms that you had more trouble making decisions about than other symptoms? If so, which ones gave you the most trouble?

Question: What sources of information did you rely on for completing this assessment?

Question: Describe any challenges that you had in obtaining the information needed to complete this assessment.

Question: Describe any challenges that you encountered that seemed to be specific to a particular setting.

Question: How easy or difficult do you think it will be to gather the information needed to complete this assessment in practice (i.e., outside of a testing situation)? What might some of the barriers be when completing this assessment in practice?

Anxiety

General comments

Question: Describe any challenges you encountered that seemed to be setting-specific.

Question: Describe any challenges in obtaining the information to complete this domain.

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Question: Did the patients express any concern about their privacy when asked about their anxiety symptoms?

Question: After you inform the patients that these questions are not meant to give them a diagnosis, but to assist planning for better care, did patients feel comfortable talking about their anxiety?

Question: Did patients have any trouble recalling their symptoms during a lookback period of 7 days?

Pain

General comments

Question: Describe any challenges you encountered that seemed to be setting-specific.

Question: Describe any challenges in obtaining the information to complete this domain.

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Medication Reconciliation

General comments

Question: Describe any challenges you encountered that seemed to be setting-specific.

Question: Describe any challenges in obtaining the information to complete this domain.

Question: Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Question: To what extent did the EHR help/hinder you in completing this section?

Med Rec–Specific Question: Was enough information available to answer? Do you think people were clear on when to answer no versus 9?

Med Rec Tablet Question: Did the format work? Built-in logic did not allow all rows to show up—how what that?

Care Preferences

General comments

Describe any challenges you encountered that seemed to be setting-specific.

Describe any challenges in obtaining the information to complete this domain.

Describe any "problem" assessment questions—what were the specific challenges associated with this question and how did you or could we resolve this issue?

Follow up Question: Did you understand what we meant by "Goals of Care" and the different areas of "Goals of Care" (i.e., physical goals, social goals, intellectual goals, emotional goals)?

To what extent did the EHR help/hinder you in completing this section?

Resources

How easy or difficult was it to use the tablet to input data? Does the way the sections in the tablet are laid out reflect your usual process?

To what extent have you encountered challenges with the tablet itself? (Probes: Tablet not turning on, not responding/freezing, crashing.)

What improvements do you suggest?

What are the best aspects of using the tablet?

How is using it different than using paper?

What are your thoughts about the crosswalk interface? How is the workflow of switching back and forth to get the ID number? Any other thoughts about the crosswalk? How can it be improved?

Did you reference the user manual for any assessment areas? If so, for which data elements and what were you looking for? In what ways did you find the user manual to be helpful to you?

What other tools could help you complete the information in the tablet?

To what extent did the training you attended prepare you for conducting the assessments and using the tablets (and beginning the pilot data-collection process)? Were the examples included in the training and in the manual helpful? Do you have any suggestions for improvements or additions regarding the examples?

General Comments on Entire Process

Do you have any other observations/thoughts about the process that you would like to share?

Action Items:

Note Taker Comments:

Changes to Interview Structure:

Additional Protocol: Medication Reconciliation Clarification

Initial feedback for the MR items led to a change in the assessment that removed F1b. In addition, RAND researchers noted some problem areas that needed more-specific clarification and drafted the below protocol for the last round of research nurse and facility/agency staff interviews. RAND research staff sat in on the interviews and conducted this portion instead of the usual interviewer.

Medication Reconciliation Additional Follow Up

For context, we have noticed in our first cut of the data (approximately 100 patients/residents) that the research nurses (RNs) and field staff (FS) appear to differ quite a bit on whether the patient/resident was on a medication(s) within a certain drug class. We wanted to ask for your input on why this might be the case.

1. Are there any thoughts you have on why there might be a mismatch between FS and RNs?
2. Why might there be more differences in noting a patient is taking a medication within a drug class, for some drug classes and less so for other drug classes?
3. How are you responding to F1c when a medication is noted as "PRN"?
4. What kinds of drug class resources are you using to determine a drug class? Do you use the PDF we provided in the tablet? Or some other resource? Have you used any resources that the RNs gave you (anything called "Drug Classification Reference by Trade Name")?
5. Are there any issues you've noted by PAC setting that would influence the difference between RNs and FS? For example, perhaps for certain settings, there is more PRN prescribing for some settings, certain FS are more likely to fill out the assessment who may have more or less familiarity with drug classes, or use of auto-populated drug classes for a facility/agency?
6. Is there any confusion between indications versus drug classes? Could you explain the difference between FS and RNs?

As a reminder, we have been asking all assessors to skip F1b. For PRN medications, you should select a "1" for F1c, days on a specific drug class, so that you are prompted to receive further items for that drug class.

References

Abdulla, Aza, Nicola Adams, Margaret Bone, Alison M. Elliott, Jean Gaffin, Derek Jones, Roger Knaggs, Denis Martin, Liz Sampson, and Pat Schofield, "Guidance on the Management of Pain in Older People," *Age and Ageing*, Vol. 42, 2013, pp. i1–57.

American Geriatrics Society Panel on the Pharmacological Management of Persistent Pain in Older Persons, "Pharmacological Management of Persistent Pain in Older Persons," *Pain Medication*, Vol. 10, No. 6, 2009, pp. 1062–1083.

Arnetz, J. E., I. Almin, K. Bergstrom, Y. Franzen, and H. Nilsson, "Active Patient Involvement in the Establishment of Physical Therapy Goals: Effects on Treatment Outcome and Quality of Care," *Advances in Physiotherapy*, Vol. 6, No. 2, 2004, pp. 50–69.

Avenanti, Alessio, Domenica Bueti, Gaspare Galati, and Salvatore M. Aglioti, "Transcranial Magnetic Stimulation Highlights the Sensorimotor Side of Empathy for Pain," *Nature Neuroscience*, Vol. 8, No. 7, 2005, pp. 955–960.

Bao, Yuhua, Huibo Shao, Tara F. Bishop, Bruce R. Schackman, and Martha L. Bruce, "Inappropriate Medication in a National Sample of U.S. Elderly Patients Receiving Home Health Care," *Journal of General Internal Medicine*, Vol. 27, No. 3, 2012, pp. 304–310.

Boockvar, Kenneth S., Heather Carlson LaCorte, Vincent Giambanco, Bella Fridman, and Albert Siu, "Medication Reconciliation for Reducing Drug-Discrepancy Adverse Events," *American Journal of Geriatric Pharmacotherapy*, Vol. 4, No. 3, 2006, pp. 236–243.

Campbell, Susan E., D. Gwyn Seymour, William R. Primrose, Joanna E. Lynch, Edmund Dunstan, Mireia Espallargues, Giovanni Lamura, Peter Lawson, Ian Philp, Elizabeth Mestheneos, Barbara Polityńska, and Ismo Raiha, "A Multi-Centre European Study of Factors Affecting the Discharge Destination of Older People Admitted to Hospital: Analysis of In-Hospital Data from the ACME*plus* Project," *Age and Ageing*, Vol. 34, No. 5, 2005, pp. 467–475.

Casten, Robin, M. Powell Lawton, Patricia A. Parmelee, and Morton H. Kleban, "Psychometric Characteristics of the Minimum Data Set I: Confirmatory Factor Analysis," *Journal of the American Geriatrics Society*, Vol. 46, No. 6, June 1998, pp. 726–735.

Centers for Medicare & Medicaid Services, "LTCHPPS Historical Impact Files," webpage, February 29, 2012. As of February 26, 2018: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/LongTermCareHospitalPPS/LTCHPPS-Historical-Impact-Files.html

———, "Inpatient Rehabilitation Facility PPS: Data Files," webpage, August 8, 2017. As of February 26, 2018: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/Data-Files.html

———, "Provider of Services Current Files," webpage, February 28, 2018. As of March 13, 2018: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/index.html

CMS—*See* Centers for Medicare & Medicaid Services.

Cohen, Jacob, "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement*, Vol. 20, No. 1, 1960, pp. 37–46.

Cotter, Valerie T., "Alzheimer's Disease: Issues and Challenges in Primary Care," *Nursing Clinics of North America*, Vol. 41, No. 1, 2006, pp. 83–93.

Data.Medicare.gov, "Home Health Compare Data Archive," webpage, undated(a). As of February 26, 2018: https://data.medicare.gov/data/archives/home-health-compare

———, "Nursing Home Compare Data Archive," webpage, undated(b). As of February 26, 2018: https://data.medicare.gov/data/archives/nursing-home-compare

Delate, Thomas, Elizabeth A. Chester, Troy W. Stubbings, and Carol A. Barnes, "Clinical Outcomes of a Home-Based Medication Reconciliation Program After Discharge from a Skilled Nursing Facility," *Pharmacotherapy*, Vol. 28, No. 4, 2008, pp. 444–452.

Graff, Maud J. L., Myrra J. M. Vernooij-Dassen, Marjolein Thijssen, Joost Dekker, Willibrord H. L. Hoefnagels, and Marcel G. M. Olde Rikkert, "Community Based Occupational Therapy for Patients with Dementia and Their Care Givers: Randomised Controlled Trial," *BMJ*, Vol. 333, No. 7580, 2006, p. 1196.

Heruti, Raphael J., Ayala Lusky, Rachel Dankner, Haim Ring, Mark Dolgopiat, Vita Barell, Shalom Levenkrohn, and Abraham Adunsky, "Rehabilitation Outcome of Elderly Patients After a First Stroke: Effect of Cognitive Status at Admission on the Functional Outcome," *Archives of Physical Medicine and Rehabilitation*, Vol. 83, No. 6, 2002, pp. 742–749.

The Joint Commission, *Comprehensive Accreditation Manual for Hospitals*, Chicago, Ill., 2015.

Kane, Robert L., and Rosalie A. Kane, "What Older People Want from Long-Term Care, and How They Can Get It," *Health Affairs*, Vol. 20, No. 6, 2001, pp. 114–127.

Kane, Rosalie A., "Goals of Home Care: Therapeutic, Compensatory, Either, or Both?" *Journal of Aging and Health*, Vol. 11, No. 3, 1999, pp. 299–321.

Landis, J. Richard, and Gary G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, Vol. 33, No. 1, 1977, pp. 159–174.

Lawton, M. Powell, Robin Casten, Patricia A. Parmelee, Kimberly Van Haitsma, Julie Corn, and Morton H. Kleban, "Psychometric Characteristics of the Minimum Data Set II: Validity," *Journal of the American Geriatrics Society*, Vol. 46, No. 6, June 1998, pp. 736–744.

Lin, Wen-Chieh, Terry Y. Lum, David R. Mehr, and Robert L. Kane, "Measuring Pain Presence and Intensity in Nursing Home Residents," *Journal of the American Medical Directors Assocation*, Vol. 7, No. 3, 2006, pp. 147–153.

Logsdon, Rebecca G., Laura E. Gibbons, Susan M. McCurry, and Linda Teri, "Assessing Quality of Life in Older Adults with Cognitive Impairment," *Psychosomatic Medicine*, Vol. 64, No. 3, 2002, pp. 510–519.

Saliba, Debra, and Joan Buchanan, *Development and Validation of a Revised Nursing Home Assessment Tool: MDS 3.0*, Baltimore, Md.: Centers for Medicare & Medicaid Services, April 2008.

Thielke, Stephen M., Joanna Sale, and M. Carrington Reid, "Identifying, Tracking, and Managing Pain in LTC," *Annals of Long-Term Care*, Vol. 18, No. 9, 2010.

Wells, Nancy, Chris Pasero, and Margo McCaffery, "Improving the Quality of Care Through Pain Assessment and Management," in R. G. Hughes, ed., *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*, Rockville, Md.: Agency for Healthcare Research and Quality, 2008.

Whitlatch, Carol J., Rich Piiparinen, and Lynn Friss Feinberg, "How Well Do Family Caregivers Know Their Relatives' Care Values and Preferences?" *Dementia*, Vol. 8, No. 2, 2009, pp. 223–243.