

Development and Evaluation of Candidate Standardized Patient Assessment Data Elements

Findings from the National Beta Test (Volume 2)

Maria Orlando Edelen, Anthony Rodriguez, Sangeeta C. Ahluwalia,
Emily K. Chen, Sarah Dalton, Michael Stephen Dunbar, Shira H. Fischer,
David J. Klein, Monique Martineau, Patrick Orr, Jessica Phillips,
Debra Saliba, Victoria Shier, Regina A. Shih



For more information on this publication, visit www.rand.org/t/RR3004

Published by the RAND Corporation, Santa Monica, Calif.

© Copyright 2019 RAND Corporation

RAND® is a registered trademark.

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of its research documents for commercial use. For information on reprint and linking permissions, please visit www.rand.org/pubs/permissions.

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

www.rand.org

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 background information on and results of the national Beta test of a set of candidate items for assessing these focus areas, conducted between November 2017 and August 2018. Volume 2 covers the candidate items tested, the design and sampling plan, information on training, recruitment, and retention, information on the data collection process, and the analytic plan. Subsequent volumes present the quantitative and qualitative data gathered during testing, as well as analytic comments and recommendations.

This work was sponsored by CMS under contract No. HHSM-500-2013-13014I. The research was carried out within the Quality Measurement and Improvement Program in RAND Health Care.

RAND Health Care, a division of the RAND Corporation, promotes healthier societies by improving health care systems in the United States and other countries. We do this by providing health care decisionmakers, practitioners, and consumers with actionable, rigorous, objective evidence to support their most complex decisions.

For more information, see www.rand.org/health-care, or contact

RAND Health Care Communications
1776 Main Street
P.O. Box 2138
Santa Monica, CA 90407-2138
(310) 393-0411, ext. 7775
RAND_Health-Care@rand.org

Research Team

Project Leadership

Maria Edelen, Ph.D. RAND Health	Project Director
Emily Chen, Ph.D. RAND Health	Project Co-Director
Susan Paddock, Ph.D. RAND Health	Lead Statistician
Anthony Rodriguez, Ph.D. RAND Health	Lead Psychometrician
Sangeeta Ahluwalia, Ph.D.	Task Lead for Training
Rosa Elena Garcia, M.P.H.	Survey Director
Debra Saliba, M.D., M.P.H.	Senior Clinical Advisor

RAND Project Staff

Emily Butcher, B.A.
Catherine Cohen, Ph.D.
Sarah Dalton, M.A.
Michael Dunbar, Ph.D.
Jason Etchegaray, Ph.D.
Shira Fischer, M.D., Ph.D.
Liisa Hiatt, M.S.
Wenjing Huang, Ph.D.
David Klein, M.S.
Jaime Madrigano, Sc.D.
Monique Martineau, M.A.
Patrick Orr, B.A.
Elizabeth Petrun-Sayers, Ph.D.
Lauren Pfeifer, B.A.
Jessica Phillips, M.S.
Francesca Pillemer, Ph.D.
Nabeel Qureshi, M.P.H.
Rachel Reid, M.D.
Rachel Ross, M.P.H.
Shoshana Shelton, M.P.H.
Cathy Sherbourne, Ph.D.
Victoria Shier, Ph.D.
Regina Shih, Ph.D.

Molly Simmons, Ph.D.
Elizabeth Sloss, Ph.D.

Abt Project Staff

Terry Moore, M.P.H., B.S.N.	Principal Associate/V.P., Health Policy
Teresa Mota, B.S.N., R.N.	Nurse Researcher/Associate
Brenda Karkos, M.S.N./M.B.A., R.N.	Associate-Nurse Researcher
Lauren Christopher	Research Assistant
Claire Hoffman	Associate Analyst
Olivia Jung	Analyst
Kelly Weiss	Research Assistant

Qualidigm Project Staff

Ann Spenard, M.S.N., R.N.-B.C.	Vice President /Principal
Lisa Newton, M.S.N., R.N.-B.C.	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

Northwestern Project Staff

Benjamin David Schalet, Ph.D.	Director, subcontract
Aaron James Kaat, Ph.D.	Psychometrician
John Devin Peipert, Ph.D.	Psychometrician
Christa Martens, M.A.	Project Manager

Contents

- Preface iii
- Acknowledgments..... viii
- Abbreviations..... ix
- Chapter 1. Introduction 1
 - Project Overview 1
 - Previous Stakeholder Input Opportunities..... 2
- Chapter 2. Candidate SPADEs Included in Beta Testing 5
 - Selection and Development of Candidate SPADEs 5
 - Cognitive Function and Mental Status 5
 - Pain 10
 - Impairments..... 11
 - Special Services, Treatments, and Interventions 12
 - Other 12
- Chapter 3. Design and Sampling 15
 - Design of Data Collection 15
 - Testing Targets 16
 - Market and Facility Sampling 18
- Chapter 4. Provider and Research Nurse Recruitment 21
 - Providers..... 21
 - Research Nurses 24
- Chapter 5. Training 26
 - Research Nurses 26
 - Facility/Agency Staff..... 27
 - Refresher Trainings 29
 - Replacement Training 29
- Chapter 6. Data Collection..... 31
 - Field Management..... 31
 - Data Management and Processing..... 33
 - Data Reporting..... 33
 - Data Delivery..... 33
 - Additional Data Collection Activities 34
- Chapter 7. Analytic Plan 37
 - Goals and Approach 37
- Appendix A. Past Performance, Current Use, and Input Opportunities for Candidate SPADEs
 - Tested in Beta 45
- Appendix B. Beta Research Nurse Training Agenda 51
- Appendix C. Example Beta Field Training Agenda (Boston) 56

Appendix D. Beta Field Staff Web-Based Survey Questions.....	60
References.....	71

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 in “Research Team.” We thank the Research Nurses and the facilities 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 Gage of George Washington University for their thoughtful reviews of the report.

Abbreviations

BIMS	Brief Interview for Mental Status
CAM	Confusion Assessment Method
CARE	Continuity Assessment Record and Evaluation
CMS	Centers for Medicare & Medicaid Services
CY	calendar year
EHR	electronic health record
FY	fiscal year
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	LTCH CARE Data Set
LTCH	long-term care hospital
MDS	Minimum Data Set
OASIS	Outcome and Assessment Information Set
PAC	post-acute care
PAC-PRD	Post-Acute Care Payment Reform Demonstration
PC	public comment
PHQ	Patient Health Questionnaire

PROMIS	Patient-Reported Outcomes Measurement Information System
SNF	skilled nursing facility
SPADE	standardized patient assessment data element
SRG	Survey Research Group
SSTIs	special services, treatments, and interventions
TEP	technical expert panel

Chapter 1. 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 data elements (SPADEs) 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 (PAI) for IRFs, LTCH Continuity Assessment Record and Evaluation (CARE) Data Set (LCDS) for LTCHs, and Minimum Data Set (MDS) for SNFs. SPADEs 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, SPADEs within the following clinical categories:

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

In consultation with CMS, we focused on identifying and evaluating candidate SPADEs from a subset of these categories in order to 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 candidate SPADE selection activities, and in consultation with CMS, RAND established the following content area work teams during the project period:

1. cognition and mental status: cognitive status
2. cognition and mental status: depressed mood
3. medical conditions: pain
4. impairments: vision and hearing; bladder and bowel continence
5. special services, treatments, and interventions
6. other: care preferences, medication reconciliation, global health.

In addition, we worked with CMS to establish 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 (then) 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.

Previous Stakeholder Input Opportunities

Candidate SPADEs 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), discussions with stakeholders, discussions with partners within CMS and the U.S. Department of Health and Human Services, and feedback from our TEP.

Prior to Beta testing, the Alpha 1 and Alpha 2 feasibility tests were conducted between August 2016 and October 2016, and between April 2017 and July 2017, respectively. Further information was collected from other stakeholder input opportunities, including two subregulatory calls for public comment, and proposed rulemaking for the fiscal year 2018/calendar year 2019 (FY 2018/CY 2019) rule cycle. See below for more detail on these activities and their corresponding reports:

- **TEP 1** (April 2016): Sixteen TEP members gathered for a two-day in-person meeting to provide input on data elements in the current PAC assessments and other identified candidate SPADEs. TEP members rated data elements on their potential for improving quality, validity, feasibility for use in PAC, and utility for describing case mix. A report of the first TEP meeting can be found at https://www.rand.org/content/dam/rand/pubs/working_papers/WR1100/WR1187/RAND_WR1187.pdf.¹

¹ Centers for Medicare & Medicaid Services (CMS), *Development and Maintenance of Post-Acute Care Cross-Setting Standardized Patient Assessment Data: Technical Expert Panel Summary/Expert Input Report*, June 2016.

- **Public Comment 1** (PC 1) (August to September 2016): A subregulatory call for public comment was solicited for a subset of the candidate SPADEs through the CMS website, for which 66 comments were received. The Public Comment 1 summary report is available at <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Development-of-Functional-Outcome-Quality-Measures-for-SNFs-Public-Comment-Summary-Report.pdf>.²
- **Alpha 1 Feasibility Testing** (August to October 2016): Alpha 1 was the first phase of pilot testing candidate SPADEs. Testing was conducted among four PAC providers (one of each PAC type) in the greater Hartford, Connecticut, area. Research Nurses and Facility Staff conducted 133 paired assessments so that results could be compared for both feasibility and interrater reliability (IRR). The Alpha 1 feasibility test report is available at <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Standardized-Patient-Assessment-Data-Element-Alpha-1-Report.pdf>.³
- **TEP 2** (January 2017): Fourteen TEP members gathered for a two-day in-person meeting to review interim results of Alpha 1 testing and other potential candidate SPADEs. TEP members rated potential candidate SPADEs on their potential for improving quality, validity, feasibility for use in PAC, and utility for describing case mix. A follow-up webinar for TEP members was held to continue the discussion of candidate SPADEs (July 2017). A report of the second TEP meeting and follow-up webinar can be found at <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/RAND-IMPACT-TEP-Second-Convening-Final-Report-March-2017.pdf>.⁴
- **Public Comment 2** (PC 2) (April to June 2017): A second subregulatory call for public comment was solicited through the CMS website, for which 33 comments were received. The Public Comment 2 summary report is available at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Public-Comment-Summary-Report_Standardized-Patient-Assessment-Data-Element-Work_PC2.pdf.⁵
- **Proposed rulemaking** (fall 2016 to summer 2017): Some candidate SPADEs that had previously undergone feasibility testing (i.e., before the Alpha 1 pilot commenced) were submitted for proposed rulemaking for the FY 2018/CY 2019 rule cycle. Results of this process are available on the Federal Register at <https://www.federalregister.gov/d/2017-07800> (for LTCHs); <https://www.federalregister.gov/d/2017-08521> (for SNFs);

² CMS, *Development of Functional Outcome Quality Measures for Skilled Nursing Facilities (SNFs): Public Comment Summary Report*, September 2016.

³ M. O. Edelen et al., *Development and Maintenance of Standardized Cross-Setting Patient Assessment Data for Post-Acute Care: Summary Report of Findings from Alpha 1 Pilot Testing*, Santa Monica, Calif.: RAND Corporation, RR-1895-CMS, March 2017.

⁴ CMS, *Development and Maintenance of Post-Acute Care Cross-Setting Standardized Patient Assessment Data: Technical Expert Panel Summary (Second Convening)*, March 2017.

⁵ CMS, *Development and Maintenance of Post-Acute Care Cross-Setting Standardized Assessment Data: Public Comment Summary Report 2*, January 2018.

<https://www.federalregister.gov/d/2017-08428> (for IRFs); and <https://www.federalregister.gov/d/2017-15825> (for HHAs).⁶

- **Alpha 2 Feasibility Testing** (April to July 2017): Testing was conducted among 15 PAC providers in three regions of the United States. As with Alpha 1 testing, Research Nurses and Facility Staff conducted paired assessment (for communicative patients, 118 at admission and 42 at discharge; for noncommunicative patients, 44 assessments) so that results could be compared for both feasibility and IRR. The Alpha 2 feasibility test report is available at https://www.rand.org/pubs/external_publications/EP67619.html.⁷

The results of these activities combined to inform the content and design of the national Beta test to evaluate candidate SPADE performance when used in any of the four PAC settings. The overarching goal of the national test is to evaluate the reliability and validity of candidate data elements and to identify the best, most feasible subset for standardization to meet requirements of the IMPACT Act.

The purpose of this report is to describe the methods and results from the Beta test. This second volume focuses on the design and methods for the national Beta test.

⁶ National Archives and Records Administration (NARA), “Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long-Term Care Hospital Prospective Payment System and Proposed Policy Changes and Fiscal Year 2018 Rates; Quality Reporting Requirements for Specific Providers; Medicare and Medicaid Electronic Health Record (EHR) Incentive Program Requirements for Eligible Hospitals, Critical Access Hospitals, and Eligible Professionals; Provider-Based Status of Indian Health Service and Tribal Facilities and Organizations; Costs Reporting and Provider Requirements; Agreement Termination Notices,” *Federal Register*, Vol. 82, No. 81, April 28, 2017, pp. 19796–20231; NARA, “Prospective Payment System and Consolidated Billing for Skilled Nursing Facilities for FY 2018, SNF Value-Based Purchasing Program, SNF Quality Reporting Program, Survey Team Composition, and Proposal To Correct the Performance Period for the NHSN HCP Influenza Vaccination Immunization Reporting Measure in the ESRD QIP for PY 2020,” *Federal Register*, Vol. 82, No. 85, May 4, 2017, pp. 21014–21100; NARA, “Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2018,” *Federal Register*, Vol. 82, No. 84, May 3, 2017, pp. 20690–20747; NARA, “CY 2018 Home Health Prospective Payment System Rate Update and Proposed CY 2019 Case-Mix Adjustment Methodology Refinements; Home Health Value-Based Purchasing Model; and Home Health Quality Reporting Requirements,” *Federal Register*, Vol. 82, No. 144, July 28, 2017, pp. 35270–35393.

⁷ M. O. Edelen et al., *Development and Maintenance of Standardized Cross-Setting Patient Assessment Data for Post-Acute Care: Summary Report of Findings from Alpha 2 Pilot Testing*, Santa Monica, Calif.: RAND Corporation, EP-67619, June 2018.

Chapter 2. Candidate SPADEs Included in Beta Testing

The Beta testing phase built on early information-gathering activities in the project and the Alpha 1 and Alpha 2 testing phases. This chapter describes the candidate SPADEs included in Beta testing, and the activities through which quantitative and qualitative feedback were gathered on them.

Selection and Development of Candidate SPADEs

Candidate SPADEs in Beta testing were identified in consultation with CMS and after a rigorous review and development process. For some candidate SPADEs, this included refinement in response to feedback obtained through stakeholder engagement and testing.

A summary table of the candidate SPADEs tested in Beta 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 input opportunities that have occurred as part of this effort leading up to the Beta test. The table also notes the data sources required to complete each candidate SPADE (i.e., patient/resident interview, chart review, observation, multiple sources).

Cognitive Function and Mental Status

Cognitive Status

Conducting cognitive assessments in PAC settings is essential to screen for cognitive impairment, rate severity of disorder, and develop a plan for care transitions. However, because cognitive status is a multidimensional construct, it may be challenging to obtain sufficient information to define the specific areas of cognitive impairment. Thus, the challenge of this category is to establish a relatively brief, standardized assessment of cognitive status that captures issues of memory, executive function, impaired communication, and cognitive skills for daily decisionmaking and safety.

Three candidate SPADEs that assess cognitive status were included in Beta testing: Brief Interview for Mental Status (BIMS), Confusion Assessment Method (CAM), and Staff Assessment of Mental Status.

Brief Interview for Mental Status (BIMS)

The Brief Interview for Mental Status (BIMS) is a performance-based (patient/resident interview) cognitive assessment that assesses repetition, recall with and without prompting, and temporal orientation. It was developed to be a brief screener to assess cognition, with a focus on learning and memory. Results of the BIMS describe cognitive status as cognitively intact,

moderate impairment, or severe impairment. The BIMS was discussed in TEP 1, and it was included in PC 1 and in the FY 2018/CY 2019 proposed rulemaking.

Confusion Assessment Method (CAM)

The CAM is an instrument that uses multiple information sources to screen for overall cognitive impairment as well as features to distinguish delirium or reversible confusion from other types of cognitive impairment. The CAM was discussed in TEP 1, and it was included in PC 1 and in the FY 2018/CY 2019 proposed rulemaking.

Staff Assessment of Mental Status

The data elements that comprise Staff Assessment of Mental Status assess long-term memory, short-term memory, memory/recall ability, and decisionmaking based on staff observation. These data elements are intended for use among patients/residents in all PAC settings who were unable to complete the interview-administered BIMS because of refusal, nonsensical answers, or inability to make him- or herself understood at least some of the time. The Staff Assessment of Mental Status was included in the Alpha 2 feasibility test and discussed in TEP 2, and it was included in PC 2.

Other Cognitive Function Data Elements

Expression and Understanding

Problems making oneself understood can be very frustrating and can contribute to social isolation and mood and behavior disorders. The inability to understand person-to-person communication can severely limit an individual's ability to associate with others and inhibit their ability to follow instructions, thereby posing a health and safety risk.

Two different versions of data elements assessing expression and understanding were included in Beta testing: (1) Speech Clarity, Makes Self Understood, and Ability to Understand Others, and (2) Expression of Ideas and Wants, and Understanding Verbal Content. In version one, clarity of patient/resident speech (Speech Clarity) is assessed as a separate data element. Makes Self Understood assesses the patient's/resident's ability to conduct social conversation in his or her primary language, whether in speech, writing, sign language, gestures, or a combination of these. Ability to Understand Others assesses comprehension of direct person-to-person communication, whether spoken, written, or in sign language or Braille. Expression of Ideas and Wants asks the assessor to consider verbal and nonverbal forms of communication and assesses whether the patient/resident can express or communicate requests, needs, and opinions, and can conduct social conversation in his or her primary language, whether in speech, writing, sign language, gestures, or a combination of these. Understanding Verbal Content is very similar to Ability to Understand Others. Expression of Ideas and Wants was included in PC 1.

Behavioral Signs and Symptoms

Behavioral disturbances, including disruptive or dangerous physical or verbal behaviors by a patient/resident directed at either themselves or caregivers, often signal distress or unmet/unrecognized needs.¹ Such disturbances put additional time and resource burdens on providers; disrupt care; result in poorer patient/resident outcomes; and place the patient/resident at risk for injury, isolation, and inactivity.² These symptoms may also disrupt the institutional or home environment and affect the safety and privacy of other patients/residents, caregivers, and staff. Exposure to aggressive behaviors can also have a negative impact on staff job satisfaction.³

The data elements that comprise Behavioral Signs and Symptoms (Presence and Frequency; Impact on Patient/Resident; Impact on Others; Rejection of Care) assess whether the patient/resident has exhibited any behavioral symptoms that may indicate cognitive impairment or other issues during the assessment period and use multiple information sources. Presence and Frequency was included in PC 1, discussed in TEP 2, included in PC 2, and tested in Alpha 2. It was also included in the FY 2018/CY 2019 proposed rulemaking.

Mental Status

Depression is the most common mental health problem in older adults, and is especially common—yet underrecognized and undertreated—in PAC settings.⁴ Undetected depression can lead to degraded physical and mental health and functioning, increased medical care utilization and costs, poor quality of life, and premature death.⁵ It can also exacerbate other chronic medical

¹ B. S. Husebo et al., “The Response of Agitated Behavior to Pain Management in Persons with Dementia,” *American Journal of Geriatric Psychiatry*, Vol. 22, No. 7, 2014, pp. 708–717; C. Ballard et al., “Management of Agitation and Aggression Associated with Alzheimer’s Disease: Controversies and Possible Solutions,” *Current Opinion in Psychiatry*, Vol. 22, No. 6, 2009, pp. 532–540.

² L. M. Dougherty et al., “Effects of Exposure to Aggressive Behavior on Job Satisfaction of Health Care Staff,” *Journal of Applied Gerontology*, Vol. 11, No. 2, 1992, pp. 160–172; A. H. Lequerica et al., “Agitation in Acquired Brain Injury: Impact on Acute Rehabilitation Therapies,” *Journal of Head Trauma Rehabilitation*, Vol. 22, No. 3, 2007, pp. 177–183.

³ Dougherty et al., 1992.

⁴ S. Banerjee and A. Macdonald, “Mental Disorder in an Elderly Home Care Population: Associations with Health and Social Service Use,” *British Journal of Psychiatry*, Vol. 168, No. 6, 1996, pp. 750–756; M. L. Bruce, G. J. McAvay, and P. J. Raue, “Major Depression in Elderly Home Health Care Patients,” *American Journal of Psychiatry*, Vol. 159, No. 8, 2002, pp. 1367–1374; K. Ell et al., “Routine PHQ-9 Depression Screening in Home Health Care: Depression Prevalence, Clinical and Treatment Characteristics, and Screening Implementation,” *Home Health Care Services Quarterly*, Vol. 24, No. 4, 2006, pp. 1–19; R. N. Jones, E. R. Marcantonio, and T. Rabinowitz, “Prevalence and Correlates of Recognized Depression in U.S. Nursing Homes,” *Journal of the American Geriatrics Society*, Vol. 51, No. 10, 2003, pp. 1404–1409; and H. Shao et al., “Diagnosed Depression Among Medicare Home Health Patients: National Prevalence Estimates and Key Characteristics,” *Psychiatric Services*, Vol. 62, No. 5, 2011, pp. 538–540.

⁵ B. D. Lebowitz et al., “Diagnosis and Treatment of Depression in Late Life: Consensus Statement Update,” *Journal of the American Medical Association*, Vol. 278, No. 14, 1997, pp. 1186–1190; G. J. Diefenbach, D. F. Tolin, and C. M. Gilliam, “Impairments in Life Quality Among Clients in Geriatric Home Care: Associations with Depressive and Anxiety Symptoms,” *International Journal of Geriatric Psychiatry*, Vol. 27, No. 8, 2012, pp. 828–835; K. Kroenke et al., “The Association of Depression and Pain with Health-Related Quality of Life,

conditions, compromise treatment participation and compliance, slow recovery from injuries and surgeries, and lead to rehospitalization.⁶ Although depression in the elderly often goes undiagnosed, prognosis is usually good when there is prompt recognition and treatment.⁷ Studies have also shown that treatment of depression in older adults can result in long-term cost savings.⁸

Four candidate SPADEs that assess mental status were included in Beta testing: Patient/Resident Mood Interview (PHQ-2 to -9), Staff Assessment of Patient/Resident Mood (PHQ-9 OV), Patient-Reported Outcomes Measurement Information System (PROMIS) Anxiety (patient/resident interview), and PROMIS Depression (patient/resident interview).

Patient/Resident Mood Interview (PHQ-2 to -9)

The nine-item Patient Health Questionnaire (PHQ-9) assesses each of the criteria for major depressive disorder outlined in the *Diagnostic and Statistical Manual of Mental Disorders (DSM)*. The PHQ-2 assesses the cardinal criteria for depression—depressed mood and anhedonia—using two items from the PHQ-9. The sensitivity of the PHQ-2, though acceptable,

Disability, and Health Care Use in Cancer Patients,” *Journal of Pain and Symptom Management*, Vol. 40, No. 3, 2010, pp. 327–341; R. L. Spitzer et al., “Health-Related Quality of Life in Primary Care Patients with Mental Disorders: Results from the PRIME-MD 1000 Study,” *Journal of the American Medical Association*, Vol. 274, No. 19, 1995, pp. 1511–1517; Y. Harris, and J. K. Cooper, “Depressive Symptoms in Older People Predict Nursing Home Admission,” *Journal of the American Geriatrics Society*, Vol. 54, No. 4, 2006, pp. 593–597; K. D. Kane, B. P. Yochim, and P. A. Lichtenberg, “Depressive Symptoms and Cognitive Impairment Predict All-Cause Mortality in Long-Term Care Residents,” *Psychology and Aging*, Vol. 25, No. 2, 2010, 446; and R. Schoevers et al., “Association of Depression and Gender with Mortality in Old Age: Results from the Amsterdam Study of the Elderly (AMSTEL),” *British Journal of Psychiatry*, Vol. 177, No. 4, 2000, pp. 336–342.

⁶ S. Ishii, J. E. Streim, and D. Saliba, “Potentially Reversible Resident Factors Associated with Rejection of Care Behaviors,” *Journal of the American Geriatrics Society*, Vol. 58, No. 9, 2010, pp. 1693–1700; S. Ishii, J. E. Streim, and D. Saliba, “A Conceptual Framework for Rejection of Care Behaviors: Review of Literature and Analysis of Role of Dementia Severity,” *Journal of the American Medical Directors Association*, Vol. 13, No. 1, 2012, pp. 11–23; M. A. Rapp et al., “Cognitive Decline in Patients with Dementia as a Function of Depression,” *American Journal of Geriatric Psychiatry*, Vol. 19, No. 4, 2011, pp. 357–363; R. C. Ziegelstein, “Depression in Patients Recovering from a Myocardial Infarction,” *Journal of the American Medical Association*, Vol. 286, No. 13, 2001, pp. 1621–1627; A. Carson et al., “The Outcome of Depressive Disorders in Neurology Patients: A Prospective Cohort Study,” *Journal of Neurology, Neurosurgery and Psychiatry*, Vol. 74, No. 7, 2003, pp. 893–896; M. R. DiMatteo, H. S. Lepper, and T. W. Croghan, “Depression Is a Risk Factor for Noncompliance with Medical Treatment: Meta-Analysis of the Effects of Anxiety and Depression on Patient Adherence,” *Archives of Internal Medicine*, Vol. 160, No. 14, 2000, pp. 2101–2107; W. Katon et al., “Behavioral and Clinical Factors Associated with Depression Among Individuals with Diabetes,” *Diabetes Care*, Vol. 27, No. 4, 2004, pp. 914–920; and J. E. Morley, “Weight Loss in the Nursing Home,” *Journal of the American Medical Directors Association*, Vol. 8, No. 4, 2007, pp. 201–204.

⁷ G. S. Alexopoulos et al., “Pharmacotherapy of Depression in Older Patients: A Summary of the Expert Consensus Guidelines,” *Journal of Psychiatric Practice*, Vol. 7, No. 6, 2001, pp. 361–376; P. A. Areán, and B. L. Cook, “Psychotherapy and Combined Psychotherapy/Pharmacotherapy for Late Life Depression,” *Biological Psychiatry*, Vol. 52, No. 3, 2002, pp. 293–303.

⁸ W. J. Katon et al., “Long-Term Effects on Medical Costs of Improving Depression Outcomes in Patients with Depression and Diabetes,” *Diabetes Care*, Vol. 31, No. 6, 2008, pp. 1155–1159; J. Unutzer et al., “Long-Term Cost Effects of Collaborative Care for Late-Life Depression,” *American Journal of Managed Care*, Vol. 14, No. 2, 2008, pp. 95–100.

is somewhat lower than that of the PHQ-9, but PHQ-2 poses a lower administrative burden. A hybrid version of the PHQ-9 (PHQ-2 to -9), in which the assessor transitions from the PHQ-2 to the PHQ-9 in cases in which a patient/resident screens positive for signs and symptoms of depression on the PHQ-2, was adopted for consideration and national testing based on stakeholder feedback. The PHQ-9 was discussed in TEP 1 and included in PC 1. The PHQ-2 to -9 was tested in Alpha 1. The PHQ-2 was included in the FY 2018/CY 2019 proposed rulemaking.

Staff Assessment of Patient/Resident Mood (PHQ-9 OV)

The Staff Assessment of Patient/Resident Mood (PHQ-9 OV) is an observational data element that can be used to identify any behaviors, signs, or symptoms of mood distress for patients/residents who cannot communicate or are unable or unwilling to participate in the PHQ-9 Patient/Resident Mood Interview. The Staff Assessment of Patient/Resident Mood was discussed in TEP 2, included in PC 2, and tested in Alpha 2.

PROMIS Anxiety

In light of the high incidence of anxiety-related distress in PAC patients/residents, a subset of self-report anxiety items from the National Institutes of Health (NIH)–supported PROMIS were identified for inclusion as possible SPADEs. The PROMIS item bank for anxiety focuses on fear (e.g., fearfulness, feelings of panic), anxious misery (e.g., worry, dread), hyperarousal (e.g., tension, nervousness, restlessness), and somatic symptoms related to arousal (e.g., cardiovascular symptoms, dizziness). Because of the length of the complete item bank and concerns that all items might not be relevant to the PAC population, we introduced a reduced list of items for consideration. PROMIS Anxiety was discussed in TEP 2, included in PC 2, and tested in Alpha 2.

PROMIS Depression

The PROMIS Depression items are also part of the National Institutes of Health Patient-Reported Outcomes Measurement Information System. They assess patient-reported negative mood, views of self (e.g., self-criticism and worthlessness), and social cognition (e.g., loneliness), as well as decreased positive affect and engagement.⁹ Because of the length of the complete item bank and concerns that all items might not be relevant to the PAC population, we introduced a reduced list of items for consideration. PROMIS Depression was discussed in a follow-up TEP webinar.

⁹ P. A. Pilkonis et al., “Item Banks for Measuring Emotional Distress from the Patient-Reported Outcomes Measurement Information System (PROMIS(R)): Depression, Anxiety, and Anger,” *Assessment*, Vol. 18, No. 3, 2011, pp. 263–283.

Pain

Pain significantly adversely affects a person's quality of life and is tightly linked to depression, social isolation, and diminished self-confidence and self-esteem.¹⁰ Pain is also associated with sleep disturbances,¹¹ functional disability,¹² and an increase in behavior problems, including agitation, irritability, and resistance to care, particularly for cognitively impaired patients.¹³ Even though pain is a common and recognizable human experience, it is often underrecognized, underdetected, and understudied among older adults.¹⁴ The current state of pain assessment is encouraging for standardization in that existing data elements have been shown to be feasible and reliable to administer across PAC settings. However, the challenges associated with pain assessment, especially among individuals with severe cognitive impairment or inability to communicate, warrant further consideration of the optimal approach for reliably assessing this domain and moving closer to the ideal state. Interview and observational candidate SPADEs to assess pain were included in Beta testing.

¹⁰ D. Amtmann et al., "Pain Affects Depression Through Anxiety, Fatigue, and Sleep in Multiple Sclerosis," *Rehabilitation Psychology*, Vol. 60, No. 1, 2015, p. 81; R. G. Hughes et al., "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*, Vol. AHRQ, Pub. No. 08-0043, Rockville, Md.: Agency for Healthcare Research and Quality, 2008; B. S. Husebo et al., "Efficacy of Pain Treatment on Mood Syndrome in Patients with Dementia: A Randomized Clinical Trial," *International Journal of Geriatric Psychiatry*, Vol. 29, No. 8, 2014, pp. 828–836; and S. Iliffe et al., "The Relationship Between Pain Intensity and Severity and Depression in Older People: Exploratory Study," *BMC Family Practice*, Vol. 10, No. 1, 2009, 54.

¹¹ J. McHugh, A.-M. Casey, B. Lawlor, "Psychosocial Correlates of Aspects of Sleep Quality in Community-Dwelling Irish Older Adults," *Aging and Mental Health*, Vol. 15, No. 6, 2011, pp. 749–755; G. Moscou-Jackson et al., "Acute Pain and Depressive Symptoms: Independent Predictors of Insomnia Symptoms Among Adults with Sickle Cell Disease," *Pain Management Nursing*, Vol. 17, No. 1, 2016, pp. 38–46.

¹² W.-C. Lin et al., "Measuring Pain Presence and Intensity in Nursing Home Residents," *Journal of the American Medical Directors Association*, Vol. 7, No. 3, 2006, pp. 147–153; A. Narayana et al., "National Breakthrough Pain Study: Prevalence, Characteristics, and Associations with Health Outcomes," *Pain*, Vol. 156, No. 2, 2015, pp. 252–259; and J. M. Teno et al., "Daily Pain That Was Excruciating at Some Time in the Previous Week: Prevalence, Characteristics, and Outcomes in Nursing Home Residents," *Journal of the American Geriatrics Society*, Vol. 52, No. 5, 2004, pp. 762–767.

¹³ Husebo, Ballard, Cohen-Mansfield, et al., 2014; Ishii et al., 2010; and S. van Almenkerk et al., "Pain Among Institutionalized Stroke Patients and its Relation to Emotional Distress and Social Engagement," *International Journal of Geriatric Psychiatry*, Vol. 30, No. 10, 2014, pp. 1023–1031.

¹⁴ A. Avenanti et al., "Transcranial Magnetic Stimulation Highlights the Sensorimotor Side of Empathy for Pain," *Nature Neuroscience*, Vol. 8, No. 7, 2005, pp. 955–960; M. P. Cadogan et al., "A Standardized Quality Assessment System to Evaluate Pain Detection and Management in the Nursing Home," *Journal of Post-Acute and Long-Term Care Medicine*, Vol. 7, No. 3, Supplement, 2006, pp. S11–S19; V. F. Engle, M. J. Graney, and A. Chan, "Accuracy and Bias of Licensed Practical Nurse and Nursing Assistant Ratings of Nursing Home Residents' Pain," *Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, Vol. 56, No. 7, 2001, pp. M405–M411; B. A. Ferrell, "Assessing Pain in the Elderly," *The Consultant Pharmacist*, Vol. 25, 2010, Suppl. A: pp. 5–10; A. Jordan et al., "The Utility of PAINAD in Assessing Pain in a UK Population with Severe Dementia," *International Journal of Geriatric Psychiatry*, Vol. 26, No. 2, 2011, pp. 118–126; and M. Smalbrugge et al., "Pain Among Nursing Home Patients in the Netherlands: Prevalence, Course, Clinical Correlates, Recognition and Analgesic Treatment—An Observational Cohort Study," *BMC Geriatrics*, Vol. 7, No. 3, 2007.

Pain Interview

Pain interview items included in Beta testing include assessment of presence of pain, frequency, severity, effect on sleep, interference with therapy- and non-therapy-related activities, and relief. Pain interview was discussed in TEP 1 and TEP 2, included in PC 1 and PC 2, and tested in Alpha 1.

Staff Assessment of Pain or Distress

Challenges associated with pain assessment, especially among individuals with severe cognitive impairment or inability to communicate, warrant further consideration of the optimal approach for reliably assessing this domain and moving closer to the ideal state. Thus, we included an observational assessment in the Beta test. Staff assessment of pain or distress was discussed in TEP 1 and TEP 2, included in PC 2, and tested in Alpha 2.

Impairments

Hearing and Vision

Hearing and vision impairments are common conditions among older adults. If unaddressed, these impairments can lead to confusion in new settings, increase isolation, contribute to mood disorders, and impede accurate assessment of other medical conditions such as cognition. Visual impairments have been associated with increased risk of falls.¹⁵ Hearing impairments can cause difficulty in communicating important information concerning medical conditions, care preferences, and care transitions. Ability to Hear and Ability to See in Adequate Light data elements (multiple information sources) are included in Beta testing, discussed in TEP 1, included in PC 1, and were proposed in FY 2018/CY 2019 rulemaking.

Bladder and Bowel Continence

Impaired bowel and bladder continence are common conditions that, if unaddressed, can affect a patient's/resident's activities of daily living, rehabilitation outcomes, skin integrity, or overall quality of life.¹⁶ Incontinence is associated with a host of negative outcomes, including sleep difficulties, inactivity, social isolation, and depression.¹⁷ Changes in continence can signal

¹⁵ R. Q. Ivers et al., "Visual Impairment and Risk of Hip Fracture," *American Journal of Epidemiology*, Vol. 152, No. 7, 2000, pp. 633–639; E. E. Freeman et al., "Visual Field Loss Increases the Risk of Falls in Older Adults: The Salisbury Eye Evaluation," *Investigative Ophthalmology and Visual Science*, Vol. 48, No. 10, 2007, pp. 4445–4450.

¹⁶ C. S. Landefeld, et al., "National Institutes of Health State-of-the-Science Conference Statement: Prevention of Fecal and Urinary Incontinence in Adults," *Annals of Internal Medicine*, Vol. 148, No. 6, 2008, pp. 449–458; E. Warshaw et al., "Clinical and Cost Effectiveness of a Cleanser Protectant Lotion for Treatment of Perineal Skin Breakdown in Low-Risk Patients with Incontinence," *Ostomy/Wound Management*, Vol. 48, No. 6, 2002, pp. 44–51.

¹⁷ I. Nygaard et al., "Urinary Incontinence and Depression in Middle-Aged United States Women," *Obstetrics and Gynecology*, Vol. 101, No. 1, 2003, pp. 149–156; Landefeld et al., 2008.

important changes in health status, making transfer of information at care transitions particularly critical. Thus, accurately assessing how often patients/residents are incontinent in order to develop management plans can enhance their clinical and functional status. Beta testing included appliance use, frequency of incontinent events (chart review), and patient/resident perceived problem or burden with bladder and bowel incontinent events (patient/resident interview). These data elements were discussed in TEP 1, tested in Alpha 1, and included in PC 2.

Special Services, Treatments, and Interventions

Special services, treatments, and interventions (SSTIs) can have a profound effect on an individual's health status, self-image, and quality of life. Assessing patients/residents for use of SSTIs in PAC settings provides important information about the severity of illness, the risk of complications and adverse health outcomes, and the intensity of resource use. Patients/residents in a PAC setting who receive SSTIs tend to use more resources than patients who do not, due to the intensity and amount of nursing care required to deliver the service, treatment, or intervention.

Nutritional approaches and other SSTIs were included in Beta testing. Nutritional approaches included parenteral/IV feeding, feeding tube, mechanically altered diet, and therapeutic diet. Other SSTIs included chemotherapy (with subitems), radiation, oxygen therapy (with subitems), suctioning (with subitems), tracheostomy care, invasive mechanical ventilator, noninvasive mechanical ventilator (with subitems), IV medications, transfusions, dialysis (with subitems), and IV access (with subitems). These data elements rely on chart review and were included in PC 1, proposed in FY 2018/CY 2019 rulemaking, and discussed in TEP 2.

Other

Care Preferences

The assessment of patient/resident care preferences and goals for care is critical to ensuring patient-centered and preference-concordant care through the course of a PAC episode and beyond. Information about patient/resident preferences and goals, used together with clinical guidelines, provides important direction for developing a care plan, selecting treatment options, and tailoring interventions. Three candidate care preferences SPADEs were included in Beta testing: importance of involvement of family/friends (interview), preferences for involvement in decisionmaking (interview), and advance directive–health care agent (chart review). These data elements were discussed in TEP 1 and TEP 2, tested in Alpha 2, and included in PC 2.

Medication Reconciliation

Approximately half of all hospital-related medication errors and 20 percent of adverse drug events (ADEs) occur during transition within, admission to, transfer to, or discharge from a hospital.¹⁸ At least 25 percent of all medication-related injuries are preventable.¹⁹ Medication reconciliation (MR), the process of obtaining a patient/resident's multiple medication lists and reconciling any discrepancies, is a cost-effective way to promote patient/resident safety by reducing errors and resulting adverse drug events. The five steps in the MR process as defined by the Joint Commission are to: (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.²⁰ A standardized set of data elements that assess MR with clear definitions of each step could better explicate processes for providers aiming to improve care, facilitate audits for assessment and adherence, and support future development of appropriate provider-level quality measures.

Medication reconciliation items included in Beta testing use multiple information sources to assess the following: classes of medications the patient/resident is currently taking; whether an indication is noted for medications; whether there were discrepancies; whether discrepancies addressed involved patient/resident or family/caregiver; whether discrepancies were communicated to a physician within 24 hours; whether recommended physician actions regarding discrepancies were carried out within 24 hours; and whether the reconciled medication list was communicated to patient/resident, prescriber, and/or pharmacy.

In addition to being tested in Alpha 1 and Alpha 2, these data elements were discussed in TEP 1 and TEP 2, and included in PC 2.

Global Health

The goal of health-related quality of life (HRQOL) assessment is to quantify, in a valid and reproducible way, the degree to which a medical condition or its treatment affects an individual's life. Assessing HRQOL through patient-reported outcomes (PROs) has the potential to improve quality of care by improving clinicians' abilities to monitor systems and treatment effectiveness, and by engaging patients in their care through better patient-physician communication. PROs

¹⁸ J. H. Barnsteiner, "Medication Reconciliation: Transfer of Medication Information Across Settings—Keeping It Free from Error," *Journal of Infusion Nursing*, Vol. 28, No. 2, Suppl., 2005, pp. 31–36; J. Rozich, and R. Roger, "Medication Safety: One Organization's Approach to the Challenge," *Journal of Clinical Outcomes Management*, Vol. 8, No. 10, 2001, pp. 27–34; and K. M. Gleason et al., "Reconciliation of Discrepancies in Medication Histories and Admission Orders of Newly Hospitalized Patients," *American Journal of Health-System Pharmacy*, Vol. 61, No. 16, 2004, pp. 1689–1695.

¹⁹ Institute of Medicine, *Preventing Medication Errors*, Washington, D.C.: The National Academies Press, 2007.

²⁰ The Joint Commission, *Comprehensive Accreditation Manual for Hospitals*, Chicago, Ill.: The Joint Commission, 2015; Y. Bao et al., "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.

include direct reports from patients/residents about their function, symptoms, and perceptions of their health and/or response to therapies. The National Institutes of Health (NIH)–supported PROMIS has developed item banks for a large number of HRQOL domains using rigorous methodology.

The Beta test included PROMIS Global Health as a candidate SPADE. It consists of ten patient/resident interview items that assess general domains of health and functioning, including physical health, mental health, social health, pain, fatigue, and overall perceived quality of life. They were included in PC 2.

Chapter 3. Design and Sampling

The national Beta test was carefully designed so that data could be collected from a wide range of environments, allowing for thorough evaluation of candidate SPADEs in all PAC settings. The test was designed to include PAC providers in a total of 14 markets across the country. This number of markets was chosen to be similar to the design used for the Post-Acute Care Payment Reform Demonstration (PAC-PRD). In addition, to meet the requirements of the IMPACT Act, it was necessary to demonstrate sufficient performance of all candidate SPADEs at admission to and discharge from each of the four PAC settings, to establish the reliability of the candidate SPADEs, and to identify whether the timing of the assessments has an influence on overall results. These requirements led to the development of a rigorous study design and sampling plan. This chapter describes the design plan for data collection and the assessment protocols, as well as the methods used to randomly sample test markets and providers throughout the United States.

Design of Data Collection

The Beta test design called for assessment data to be collected by trained Research Nurses, as well as trained Facility/Agency Staff from each participating provider. The role of the Research Nurse was to oversee field data collection in their market and to serve as a “gold standard” coder alongside a Facility/Agency Staff person, allowing measurement of IRR. This approach brought necessary robustness and consistency to both the testing and training processes. (For more information on training, see Chapter 5, “Training.”)

To test performance of candidate SPADEs for different types of patients/residents, as well as the effect of timing of administration on results, we set up assessment protocols as described below.

Assessment Protocols

Testing included three types of protocols for candidate SPADEs (see Table 3.1):

1. **Communicative admission assessment**, to be administered upon admission to a PAC site among communicative patients/residents who could make themselves understood using any means (i.e., writing, gesturing, speaking)
2. **Communicative discharge assessment**, to be administered at or near discharge from a PAC site among communicative patients/residents who could make themselves understood using any means (i.e., writing, gesturing, speaking)
3. **Noncommunicative assessment**, to be administered at any point during a qualifying PAC stay among noncommunicative patients/residents who were unable to make themselves understood in any fashion.

Table 3.1. Beta Protocols

Protocol	Patient/Resident Eligibility	Assessment Window^a
Communicative admission	Admitted to PAC site under a new Medicare Part A or Part C covered stay; able to make themselves understood using any means	Day 1 (admission date) through day 7
Communicative discharge	Any patient/resident for whom a Beta test admission assessment had been conducted	Day 1 (discharge date) through day 3 (two days prior to discharge date)
Noncommunicative	Any patient/resident who is unable to make themselves understood using any means	48 hours, beginning on any day during the PAC stay the Research Nurse and field staff partner identify

^a Assessment window refers to the time frame during which the assessment must be conducted.

Two distinct forms were created for the communicative admission and discharge assessment protocols in order to test minor variations of five of the candidate SPADEs (expression and understanding, pain interview, and PROMIS Global Health, Anxiety, and Depression). For example, the pain interview SPADEs were tested using “in the past 3 days . . .” in one form and “in the past 5 days . . .” in the second form. The 14 Beta markets were stratified by region and randomly split so that assessors from half the providers completed one testing form and the other half completed the other testing form. The complete protocols for Beta are available online,¹ and the form variations are indicated in the protocols.

Testing Targets

The target numbers of providers by setting were 28 IRFs, 28 LTCHs, 84 SNFs, and 70 HHAs, totaling 210 PAC providers. The rationale for these numbers is provided in the market and facility sampling section below. Patients and residents who were receiving care at one of the participating provider sites and were Medicare beneficiaries covered under one of the PAC prospective payment systems were eligible for inclusion. The target number of assessments per facility/agency is shown in Table 3.2.

¹ Communicative admission: Centers for Medicare & Medicaid Services, “National Field Test Assessment Protocol: Admission,” November 2017; communicative discharge: Centers for Medicare & Medicaid Services, “National Field Test Assessment Protocol: Discharge,” November 2017; and noncommunicative: Centers for Medicare & Medicaid Services, “National Field Test Assessment Protocol: Non-Communicative,” November 2017.

Table 3.2. Target Sample Sizes for Beta Test Assessments for Communicative Admission, Communicative Discharge, and Noncommunicative, by PAC Setting

Assessment Type	LTCH (n = 28)	IRF (n = 28)	SNF (n = 84)	HHA (n = 70)	Total (n = 210)
Communicative admission	840	840	2,100	1,750	5,530
Communicative discharge	579	756	1,491	840	3,666
Noncommunicative	280	280	420	350	1,330

The Beta test design aimed to collect communicative admission assessments from 25 to 30 patients/residents at each facility/agency, yielding a targeted total of 5,530 communicative admission assessments across the four PAC settings (target of 840 in LTCH, 840 in IRF, 2,100 in SNF, and 1,750 in HHA). A subset of the admission assessments (target of ten from each LTCH and IRF, five from each SNF and HHA) were conducted as paired assessments, administered by both Research Nurses and Facility/Agency Staff at admission. Data from the paired assessments were used to determine IRR. Another subset of the admission assessments (target of five from each LTCH and IRF, two from each SNF and HHA) were conducted as repeat assessments. This subset of patients/residents was assessed initially on admission day three, and then assessed again on a subset of the candidate SPADEs on admission day five and admission day seven. The repeat assessments were conducted to evaluate the impact of different assessment dates on SPADE performance. The candidate SPADEs included in the repeat assessment are identified in the protocols that are available online.

Facility/Agency Staff or Research Nurses were also enlisted to conduct discharge assessments on all communicative patients/residents who completed an admission assessment and were subsequently discharged from their PAC facility during the data collection period. These discharge assessments are considered “matched” to the admission assessment of the same patient/resident. The available number of patient/resident discharges was expected to vary by PAC setting, based on patient flow and typical duration of stay. Targeted completion of discharge assessments during Beta testing was estimated accordingly.

For noncommunicative patients/residents, both Facility Staff and Research Nurses administered the noncommunicative protocol, which constitutes three candidate SPADEs developed specifically for use with patients/residents who are unable to communicate. Reliability for these candidate SPADEs was assessed using the Facility/Agency Staff and Research Nurse assessment pairs for the noncommunicative protocol.

The completed assessment targets were set to ensure sufficient data to complete analyses of outcomes, validity, reliability, and look-back periods. Specifically, the sample size targets were selected to provide 80-percent power to detect moderate to small standardized differences: of at least a Cohen’s $d = 0.14$ – 0.38 standard deviation between any two settings for the validity

analyses; to detect differences between kappa = 0.4 (a relatively low value) versus 0.7 for a substantial range of percentage values for the reliability analyses; and to detect setting-level differences between assessment points (admission day three, admission day five, and admission day seven) corresponding to a Cohen's $d = 0.19$ standard deviation. Power calculations were based on comparing data from the LTCHs and IRFs because we have fewer of these two setting types. Thus, power to compare data from SNFs and HHAs under this design would be greater than described above.

Routine assessment data (from OASIS-C2, IRF-PAI, LCDS, and MDS 3.0) were also obtained from CMS for all patients/residents who participated in Beta testing. Data from these routine assessments were used to contribute to the evaluation of the validity of the candidate SPADEs. See Chapter 7, "Analytic Plan," for more detail.

Market and Facility Sampling

A multistage stratified random sampling plan was used to obtain the sample of 14 geographic/metropolitan areas, or "markets," in the United States, and then a sample of eligible PAC facilities was compiled from those markets. The sampling plan described below was chosen to reach the targeted number of providers while accommodating the practical need to concentrate Research Nurses in geographic areas.

Stage 1: Defining the Markets

For the first stage of sampling, RAND statisticians stratified the U.S. geographic regions by nine census divisions to ensure representativeness across the country. Hospital referral regions (HRRs), regional health care markets for tertiary medical care that generally require the services of a major referral center, were used to define the Beta test markets. There are 306 HRRs that cover the entire country and include rural and urban areas, thus offering a way to develop a nationwide probability sample of markets. RAND made the determination that at least one HRR per census division would be selected to ensure geographic diversity of the Beta sample. The Denver, Colorado, HRR was eliminated from the Beta sampling frame because of the difficulty experienced during Alpha 2 in recruiting IRFs and LTCHs for this region, as well as the market's relatively small numbers of such facilities, leaving 305 eligible HRRs. Since many markets have few or no LTCHs, RAND retained all markets as eligible provided they had at least 12 SNFs, at least ten HHAs, at least four LTCHs *or* IRFs, *and* at least one LTCH. Applying this exclusion resulted in consideration of 73 markets.

The next step in defining the sample frame was to limit the set of eligible markets to those that would have the targeted number of facilities within driving distance of each other. For each market, we determined how many eligible facilities of each PAC type were within two hours of one another, and eliminated all facilities that did not meet this requirement. This resulted in 64 eligible markets. From these eligible markets, with selection probabilities proportional to the

total number of eligible PAC facilities per market, one market was randomly sampled from each of the nine U.S. census regions to allow for regional variation in market selection, and then five additional facilities were randomly sampled irrespective of census region, resulting in the following 14 markets:

- Boston
- Chicago
- Dallas
- Durham
- Fort Lauderdale
- Harrisburg
- Houston
- Kansas City
- Los Angeles
- Nashville
- Philadelphia
- Phoenix
- St. Louis
- San Diego.

Stage 2: Identifying Providers for Sampling

Within each market, providers of each PAC type were selected from Medicare administrative files. All providers were selected from lists that were obtained from the CMS Provider of Services (POS) file (data version December 2016). The POS file contains a record for each Medicare-approved provider and is updated quarterly. In addition, LTCH providers needed to be included in the FY 2017 LTCH Final Rule Impact File, which contains data for each LTCH that was used to estimate policy updates to the LTCH prospective payment system for FY 2017. Similarly, IRF providers needed to be listed in the FY 2017 IRF Final Rule Impact File, and SNF and HHA providers needed to be listed in CMS's utilization and payment public use files released in 2017, with data covering calendar year 2014. These files included number of Medicare discharges per year, which was important for identifying facilities of sufficiently large size to yield the minimum number of Beta assessments. To be eligible, providers were required to have at least 72 Medicare discharges annually (for LTCHs and IRFs) and 60 stays annually per SNF and 60 episodes for HHA. The size requirement for LTCHs and IRFs was higher because the design calls for more patient assessments per facility in these settings than in SNFs and HHAs.

To avoid involving facility nurses who were trained as part of Alpha 1 or Alpha 2, PAC facilities that participated in those evaluations were deemed ineligible to be sampled for the Beta test. RAND also prioritized PAC facilities that enabled efficiency of data collection by giving preference to PAC facilities with drive times not exceeding two hours within each HRR, even though this resulted in lower than nationally representative percentages of rural facilities. Due to

their relatively low numbers, all eligible LTCHs and IRFs in the 14 markets were included in the initial sample of facilities for recruitment. This initial sample also targeted selection of five times as many SNF facilities and HHA agencies as were required in each market (30 SNFs, 25 HHAs) to allow for the expectation that a large proportion of invited facilities would decline participation. Overall, larger markets contributed more than the targeted number of facilities/agencies, which offset some of the smaller markets that did not have sufficient numbers of facilities/agencies to meet these targets. Additional facilities/agencies were selected for recruitment on an as-needed basis throughout the recruitment phase to meet the target numbers of settings in each market.

Recruitment procedures are described in full detail in the next chapter.

Chapter 4. Provider and Research Nurse Recruitment

With the plans for the testing design and sampling in place, the next step in setting up the Beta test was to recruit providers to participate in the study and Research Nurses to administer the assessments. This chapter describes efforts to recruit both providers and Research Nurses.

Providers

As described in the previous chapter, we sought to recruit 210 PAC providers in 14 markets across the United States. Primary recruitment activities began in March 2017 and ended in December 2017, with ongoing retention efforts taking place throughout the data collection phase. The recruitment team consisted of staff from Abt Associates and RAND's Survey Research Group (SRG).

Initial Outreach and Material Preparation (March–May 2017)

RAND and Abt Associates kicked off outreach efforts at the Medicare Learning Network (MLN) presentation on March 29, 2017, followed by an in-person stakeholder meeting held at CMS's offices in Baltimore, Maryland. Meeting attendees included representatives from major PAC corporations (i.e., Amedisys, Bayada, Healthsouth/Encompass, Kindred, Partners Healthcare/Spaulding, Select Medical Corporation, and Signature Healthcare), as well as from industry stakeholders. Stakeholder groups in attendance included the National Association for Home Care & Hospice, the Visiting Nurse Association of America, the American Medical Rehabilitation Providers Association, the National Association of Long-Term Hospitals, the American Health Care Association, the National Association for the Support of LTC, the American Hospital Association, and the California Hospital Association.

Following this meeting, the recruitment team conducted telephone outreach to corporate contacts for the major PAC corporations and stakeholders cited above and hosted targeted informational webinars for sampled facilities. The recruitment team prepared an informational postcard, a dedicated website hosted by RAND.org, a webpage on CMS.gov, and other print and electronic materials (e.g., flyers, fact sheets) to distribute directly to eligible providers and for provider associations to circulate to their members. The team also established a toll-free hotline, an email inbox, and triage procedures for addressing inquiries/expressions of interest in a timely manner. These outreach activities provided information about the Beta test goals and structure, data collection requirements, the vital role of participating providers, incentives for participation, how the data from the Beta test would be used to inform federal decisionmaking, and opportunities for questions and answers.

RAND and Abt identified, tested, and purchased licenses for customer relations management (CRM) software to help track providers through the recruitment phase, manage communications, and provide reports on the status of recruitment by region, provider type, and staff recruiter. The team worked to obtain accurate identifying information (e.g., contact name, telephone number, email address) for key decisionmakers at each facility/agency in the sample through web searches, cold calls, and acquired stakeholder contact lists.

Active Recruitment, Phase 1 (Late May–October 2017)

Active recruitment began in late May 2017. Eligible providers first received a mailed informational postcard, followed shortly by recruitment packages, which contained an invitation letter from the project team, a message from CMS, a project fact sheet, and the participation agreement. Recruiters followed the mailings with telephone calls to the eligible facilities/agencies. All recruiters used a standard call script and process. The process was as follows:

- Call the point of contact (generally the Administrator or Director of Nursing).
- Walk through the telephone recruitment script if successful in reaching the contact.
- Leave a voicemail or message with the organization’s secretary (a.k.a. the gatekeeper) with a request for a call back if unsuccessful in the first attempt.
- Document the attempt or the contact in the CRM software.
- Repeat the process every other day for up to three attempts.

In discussions with potential recruits, recruiters introduced themselves, identified the project as a CMS-funded project, inquired as to whether the facility or agency had received the recruitment mailing (and emailed the recruitment materials during or immediately following the call if the contact claimed not to have received it), and discussed the purpose of the project, the data collection requirements, and reasons why the recruit should consider participating. In some cases, a clinical staff member was enlisted to talk with the potential recruit to have a more in-depth discussion of how the facility/agency might manage the additional assessment workload, or to answer questions about the candidate SPADEs to be tested.

Once a recruited provider agreed to participate, the participation agreement was emailed for signature; follow-up calls and email reminders were sent to elicit return of the signed agreement. Upon receipt of the signed agreement, the recruitment team introduced the provider to the training team for continued engagement. The training team then worked to ensure that the recruited provider identified two Facility/Agency Staff to be data collectors for the Beta test, engaged the recruited provider as needed until the training scheduled for their market, and held the training.

Recruiters assigned a status of “unresponsive” in the tracking software to facilities/agencies that failed to call back after repeated attempts.

Active Recruitment, Phase 2 (October 2017–January 2018)

To ensure a sufficient number of providers in the Beta test, a second phase of active recruitment was added, along with a second round of in-person provider trainings in late January 2018. The second round of trainings allowed the study team to accommodate the Phase 2 provider recruits and to have an additional training date option for Phase 1 recruits that had encountered scheduling conflicts.

The second phase of active recruitment began in October 2017 and followed the same process as Phase 1. Postcards were distributed, recruitment packages were delivered, and telephone calls were made.

As in Phase 1, once a recruited provider agreed to participate and submitted a signed letter of agreement, the recruitment team introduced the provider to the training team for continued engagement. The training team ensured that the recruited provider identified two Facility/Agency Staff to be data collectors for the Beta test, engaged the recruited provider as needed until the training scheduled for their market, and held the training.

Attrition

Some providers who had agreed to participate in the Beta test ultimately ended up withdrawing from the study. Generally, providers withdrew at this stage of the study due to changes in management or administration during the interval between sign-up and training, staffing concerns, or being unavailable to attend the in-person training. Table 4.1 displays, by market, the total number of providers that agreed to participate in the Beta test when contacted for recruitment, as well as the subset of providers who participated in training.

Table 4.1. Number of Facilities/Agencies Recruited and Subset That Participated in Training by Market

Market	Recruited	Participated (% of recruited)
Boston	24	20 (83%)
Chicago	15	10 (67%)
Dallas	18	12 (67%)
Durham	16	13 (81%)
Fort Lauderdale	13	13 (100%)
Harrisburg	16	9 (56%)
Houston	17	16 (94%)
Kansas City	16	13 (81%)
Los Angeles	22	17 (77%)
Nashville	18	14 (78%)
Philadelphia	16	14 (88%)

Market	Recruited	Participated (% of recruited)
Phoenix	14	11 (79%)
San Diego	19	10 (53%)
St. Louis	16	14 (88%)
TOTAL	240	186

Table 4.2. Number of Facilities/Agencies That Participated in Training by PAC Type

PAC Type	Participated
HHA	57
IRF	28
LTCH	28
SNF	73

Research Nurses

The Beta test design planned for each of the 14 markets to be supported by two Research Nurses who reside in the area (for a total of 28). Research Nurses were assigned to sites in their market and worked in partnership with the Facility/Agency Staff at their sites to collect data. In addition to collecting assessment data, the Research Nurses also (1) trained and mentored Facility/Agency Staff partners on the candidate SPADE and data collection process; (2) assigned study IDs and confirmed patient/resident study eligibility with their Facility/Agency Staff partners; (3) tracked and monitored progress toward data collection goals; and (4) addressed data collection issues and conducted troubleshooting throughout the data collection process. Research Nurses also served as the primary liaison between the participating providers and the project team, ensuring information flow among all parties. Each Research Nurse in a market was given primary responsibility for approximately half the recruited provider sites in that market, though both Research Nurses supported each other's data collection as needed. Research Nurses were unaffiliated with the participating PAC providers.

Recruitment Process

Research Nurse recruitment occurred between June and September of 2017. Qualidigm led activities for interviewing, hiring, and managing the Research Nurses, with support from their partner agency, Ready Nurse, on recruitment, candidate prescreening, and phone interviews. Recruitment criteria included geographic proximity to facility/agency sites, experience conducting assessments in one of four PAC settings, familiarity with electronic health record (EHR) systems and technology, interest in research and health care improvement, and experience working collaboratively with patients, families, and provider staff; management experience was preferred. Successful candidates were referred to Qualidigm staff, who used a standard interview

guide with candidates to ensure comparability across interviews. All candidates were interviewed twice, with an additional interview added if any discrepancy occurred between the first two interviewers. Research Nurses were successfully recruited for all markets prior to the start of Research Nurse training.

Attrition and Replacement Hiring

Some recruited Research Nurses did not participate for the duration of the Beta test, for various reasons that are more fully described in Chapter 5. In these cases, replacement recruitment and hiring was initiated so that data collection could continue with little disruption. Replacement hiring followed the same process described above, with additional market insight provided by existing Research Nurses regarding facility/agency geography, optimal Research Nurse location, and possible need for facility reassignment.

Chapter 5. Training

Data collection during Beta was done electronically using handheld tablets, as had been pilot tested during Alpha 2. Electronic data collection facilitated the collection of data in a timely and efficient manner across hundreds of facilities and data collectors; allowed frequent transfer of data electronically from the field to RAND, ultimately improving data security; and made possible timely updates and corrections during the data cleaning process.

To ensure rigorous data collection during the Beta test, we implemented a comprehensive multicomponent training plan, consisting of virtual and in-person Research Nurse training and 18 market-specific (i.e., local) in-person Facility/Agency Staff trainings. The first Beta training phase ran from September 2017 through January 2018 to ensure that all Research Nurses and approximately 350 Facility/Agency Staff nationwide were adequately trained and prepared to collect data at their sites. A second Beta training phase for newly identified or replacement Research Nurses or Facility/Agency Staff, as well as refresher trainings and practice data collection, ran from December 2017 through March 2018. Each component of the beta training plan is described in detail below.

Research Nurses

The Research Nurse training was divided into four key components:

- e-modules providing initial training content to be completed prior to attending the in-person component
- a pretraining webinar that provided relevant study context and prepared nurses for the week of in-person training activities
- a week-long in-person training on the administration of the candidate SPADEs and hands-on practice with the assessment tool and electronic data collection procedures
- a webinar on repeat assessment data collection.

E-Modules

Three training e-modules were developed to orient Research Nurses to the project before the in-person training, to reduce training content during in-person events, and to offer on-demand/virtual options. These modules included a project overview and an overview of the national Beta test; training on confidentiality procedures, data safeguarding, and scientific integrity; and tips and techniques for conducting the Beta assessments.

Pretraining Webinar

This two-hour webinar, held the week before the in-person training, provided essential context for the study, for the Beta test, and for the Research Nurses' role as trainers and "gold standard" data collectors. Specific topics included: an overview of the Beta test phase in the context of the larger IMPACT study, including the structure and timeline, the role of the nurse and field staff trainings, and the data collection phase and purpose; an orientation to the Beta assessment form and accompanying user manual, for offline study prior to the in-person training; and moderated Q and A and review of training logistics.

In-Person Research Nurse Training

The in-person Beta Research Nurse training was held from Monday, October 2, through Friday, October 7, 2017, at the RAND offices in Santa Monica, California. We provided the 28 Research Nurses detailed instruction on the background and administration of each of the proposed Beta SPADEs. Research Nurses gained extensive hands-on practice with the assessment tool and tablet functionality via role-play/modeling and skills-based exercises. In addition, a single day of training was dedicated to preparing the nurses for field staff training, including a train-the-trainer presentation and review of presentation tips and time to practice giving assigned training presentations and receiving peer and expert feedback. During the training, Research Nurses were given all materials necessary for data collection (e.g., tablets, user manuals, paper forms, tracking sheets, secure storage bins, clipboards, cue cards, etc.). See Appendix B for the Research Nurse training agenda.

Repeat Assessment Training

In addition to the in-person Research Nurse training, RAND held a webinar in February 2018 to train Research Nurses on data collection for repeat assessments. Initiation of the repeat assessment data collection was delayed to allow Research Nurses to participate in field staff trainings through fall 2017 and gain solid experience and comfort with conducting the standard Beta assessment. The repeat assessment training webinar covered: an overview of the repeat assessment design; the process for collecting repeat assessment data, including how to schedule the assessment dates and communicate the process to patients/residents; and the procedure for how to track completion of repeat assessments and use the tablet to collect data, including a "live" walk-through on a tablet.

Facility/Agency Staff

Facility/Agency Staff who were identified by their participating Beta providers and had agreed to participate in data collection (on average, two per participating provider) were required to attend an in-person training held over two days at a central location in their market prior to beginning assessments. In-person training was conducted by project staff from RAND and

Qualidigm, with assistance from the recently trained Research Nurses who would be working in the markets. These in-person trainings allowed Research Nurses to build rapport with the staff with whom they would work, staff to ask questions of experienced trainers, staff to practice conducting aspects of the Beta assessment with feedback from trainers, and the training team to troubleshoot any issues associated with data collection tablets.

While two full days of training is a substantial commitment, this amount of time was determined to be necessary to adequately cover using the tablet to collect data, data collection procedures, and SPADE orientation. The field trainings were approved Continuing Nursing Education activities through the Connecticut Nurses’ Association, an accredited approver by the American Nurses Credentialing Center’s Commission on Accreditation. Facility/Agency Staff who attended the entire two-day training and submitted an evaluation form were awarded 14 contact hours that could be used for continuing education credit. Facility/Agency Staff training content mirrored that used for Research Nurse training; see Appendix C for the field staff training agenda.

Following the in-person training, Facility/Agency Staff were asked to conduct practice assessments at their site to augment their training within the unique context of their site, with guidance and oversight from their Research Nurse partner. The total number of practice assessments completed by staff varied based on training needs; at least one and as many as six practice assessments were conducted by Facility/Agency Staff before starting active data collection.

Two phases of training were conducted to reflect the two waves of provider recruitment; Phase 1 was held between October and November of 2017, and Phase 2 was held in January 2018. The second phase trained staff in the San Diego, Fort Lauderdale, and Chicago markets, as well as staff in other markets who may have replaced staff trained during Phase 1. The first served as test runs to smooth out on-the-ground logistics, identify challenges, and finalize field training materials. After these first trainings, RAND held a teleconference with all Research Nurses that same week to share lessons learned across the training teams. Subsequently, staff trainings were held simultaneously across markets during each training week (see Table 5.1).

Table 5.1. Facility/Agency Staff Training Schedule

Phase	Training Dates	West Region	Central Region	East Region
Phase 1	October 9 and 10	Los Angeles		
	October 18 and 19			Boston
	October 23 and 24	Los Angeles		
	October 25 and 26	San Diego	Kansas City	Philadelphia
	November 8 and 9	Phoenix	St. Louis	Harrisburg
	November 15 and 16	Dallas	Nashville	Durham
	November 29 and 30	Houston	Chicago	Ft. Lauderdale
Phase 2	January 30 and 31	San Diego	Chicago	Ft. Lauderdale

Refresher Trainings

Because of the lengthy data collection field period (i.e., approximately six months with an extended period June 1 to August 15, 2018), and because of the potential time lag between when data collectors were trained and when they began to actively collect data, we made a concentrated effort to provide refresher training to both Research Nurses and Facility/Agency Staff:

- **Refresher Training Webinars (January 11, 16, 18, 2018):** Refresher webinars were targeted primarily to Facility/Agency Staff who had been trained in fall 2017 and were just beginning to collect data for the project. Research Nurses also attended these webinars. Content included Beta test target goals; step-by-step instructions for the data collection process; a “live” walk-through of how to use the tablet to track assessments and collect data; and a review of the chart review–based candidate SPADEs, including continence and SSTIs. The agenda also included time for Q and A and sharing of resources and project assistance available to data collectors.
- **Repeat Assessment Refresher Webinar (March 29, 2018):** Because we initiated data collection for repeat assessments (i.e., day three, day five, and day seven assessments) later in the overall data collection, we held a refresher webinar for Research Nurses focused only on the repeat assessment data collection process. Content covered included a review of the repeat assessment design (e.g., eligibility rules, candidate SPADEs, goal numbers), how to track completion of repeat assessments, and a “live” walk-through of a practice scenario using the tablet assessment form.
- **Weekly Refreshers on Research Nurse Teleconference:** We held weekly check-in calls with Research Nurses throughout the data collection phase to encourage sharing of data collection progress and strategies. Time was reserved during these weekly calls to address a specific training-related topic, typically generated by project staff based on an informal assessment of current training needs, but sometimes generated by questions raised by Research Nurses and Facility/Agency Staff.

Replacement Training

Over the course of data collection, there were 11 Research Nurse resignations. Five nurses left for new jobs, three left due to illness, two were dismissed for performance issues, and one after misunderstanding job expectations. Qualidigm re-recruited and replaced Research Nurses lost due to attrition. Newly hired replacement Research Nurses were trained in a format best suited to their unique needs, following a structured “rescue” curriculum. If possible, Research Nurses attended the two-day Facility/Agency Staff training program held at a location closest to them with additional one-on-one training provided by another Research Nurse to emphasize role-specific responsibilities and content. If attending a scheduled program was not feasible, or if the Research Nurse joined the project after February 1, 2018, a two-day, one-on-one educational session was provided. The decision regarding the approach was based on maintaining standard content, cost-effectiveness, scheduling, and expedient training-to-field deployment.

A similar rescue curriculum was implemented when Facility/Agency Staff were replaced during the data collection period. The curriculum consisted of six hours of training that could be customized in two-hour blocks, if needed, and it focused on understanding the SPADEs, patient/resident interviewing techniques, and use of the tablet to collect data.

Chapter 6. Data Collection

Beta data collection began immediately after training and practice assessments were complete. Because Facility Staff trainings were staggered by market, the start of data collection was also staggered, beginning anywhere from November 2017 through February 2018. Data collection was expected to run for six months, starting from the time a site's Research Nurses and Facility/Agency Staff had completed training and ending across all markets by May 31, 2018. However, given delays in starting Beta data collection associated with the practice assessment phase and with the winter holiday season, data collection was extended through August 15, 2018.

During data collection, the two (or three) Research Nurses in each market were assigned oversight of specific PAC providers, typically from four to seven providers each. Within these sites, Research Nurses worked directly with the participating Facility/Agency Staff to collect assessment data. As described in Table 3.1, assessments could be conducted solo (unpaired) by either the Research Nurse or the field staff for communicative admission or discharge assessments, by the Research Nurse only in the case of repeat assessments, or together (paired) by both the Research Nurse and the field staff. Windows for data collection varied across assessment types (e.g., admission assessments had a seven-day window beginning on the date of admission, while discharge assessments had a three-day window ending on the date of discharge), requiring careful tracking by data collectors.

Field Management

Due to the complex data collection design and procedures, as well as the lengthy data collection time frame, project leaders deployed various mechanisms to support Research Nurses and Facility/Agency Staff during data collection, described below.

Market Manager Model

To optimize data collection and provide effective management of resources in the field, we created a system to supervise data collection activities in each of the 14 Beta markets. This model consisted of a market manager assigned to two to three markets, a regional manager assigned to one to two regions (approximately seven markets each), and field supervisors from both Qualidigm and RAND to provide project leadership support as needed. The role of the market managers was to track and report weekly data collection progress in their assigned markets, and to capture and triage issues arising in their markets as appropriate. The role of the regional managers was to serve as a nurse liaison for mentoring, troubleshooting, resolving issues, and networking among the Research Nurses. Weekly market-specific calls were

established for the market manager, regional manager, and Research Nurses to review data collection progress and challenges for each site. The market manager maintained a record of each call and reviewed the weekly assessment progress reports during the meeting. Group discussion and collective problem-solving occurred for sites that were not engaging or were experiencing challenges to produce assessments. Problems that could not be adequately addressed by those attending the call were escalated to the field supervisors for resolution. In addition to the weekly market calls, all market managers held a weekly internal call with project leadership, to provide the managers with support in their role.

Data Collection Assistance

RAND developed and maintained a project website for the Beta test, which housed training materials, webinar slide decks, and tip sheets for data collectors, as well as paper copies of the assessment forms and other project materials such as scripts/talking points, project information sheets, and cue cards. Data collectors could reach RAND project staff for help with data collection by submitting a form on this website, by sending an email to impactbetahelp@rand.org (dedicated to the Beta test), or by directly emailing or calling the training directors (Ahluwalia/Etchegaray) and/or project director (Edelen). All questions were addressed within 24 hours of receipt.

RAND also established a technical support pathway for any issues related to the tablet arising in the field, consisting of a toll-free number as well as a website where one could report problems and receive a response from technical support staff within 24 hours of receipt.

Provider Engagement and Retention Activities

Given the lengthy data collection period, priority was placed on ensuring ongoing engagement and retention of Beta test providers. To this end, several activities were implemented:

- *Newsletter*: A monthly project newsletter describing data collection progress, tips and hints for data collectors, updates regarding data collection, and upcoming CMS IMPACT-related webinars was sent to all Beta test providers.
- *Provider webinars*: We held a series of webinars during the data collection phase to emphasize the role and importance of Beta test providers, to share tips for successful data collection, to encourage brainstorming among providers regarding best practices in data submission, and to share project announcements.
- *Outreach to corporate leadership*: At the outset of data collection, we engaged with corporate leadership of PAC providers with multiple sites (Encompass, Ensign, and Kindred) participating in the Beta test to gauge their perceptions of the progress of data collection, identify potential barriers or challenges to data collection, and solicit their assistance with increasing data collection where needed.

Data Management and Processing

After data collection began at each site, staff were asked to submit completed assessments in a timely manner to allow the RAND SRG team to receive data submissions and begin quality assurance (QA) of all downloaded assessments. This multistep QA process first entailed verifying that all key variables that tracked site, survey type (Research Nurse or Facility/Agency Staff assessment), assessment type (admission communicative, noncommunicative assessment, or discharge), and interrater status (whether it was paired or solo) were complete. This last variable was critical because only a subset of admission assessments were completed by paired raters and thus needed to be tracked separately. For assessments identified with missing or incomplete variables, the SRG team made appropriate corrections or recoded information where possible. All other assessments were referred to the appropriate market manager and Research Nurse for more detailed investigation. Once a resolution was identified, SRG staff made the appropriate corrections to the submitted assessment as well as to the raw data file. Throughout, SRG staff maintained a problem log to track identified data/coding issues and their resolution.

Data Reporting

RAND SRG generated several progress reports each week based on the assessments that had been submitted:

- *Market Report*: This report displayed information for all completed assessments that were submitted for each type of assessment. It listed each of the sites within the market and their corresponding PAC type, and then included counts of each assessment type submitted, based on specific criteria. For example, paired assessments were counted and included in the report only when *both* the Facility Staff and Research Nurse assessments had been submitted and received. Discharge assessments were counted and included in the report only when the corresponding admission assessment had been received. Weekly reports were cumulative.
- *Research Nurse Report*: This report tracked all the completed nonpractice, practice, and incomplete assessments, as well as repeat assessments submitted by each Research Nurse. Weekly reports were cumulative.

As part of the QA procedures, the SRG team kept other internal logs and reports that it used during the data submission process to keep track of issues or problems with submissions.

Data Delivery

RAND SRG periodically delivered to the analytic team a cumulative file of all submitted assessments received. In addition to the assessment data variables, the SRG team provided process-level variables that included: linking ID, status of each record (finalized, pending), and notes for the analytic team (e.g., duplicate case, problem case). The analytic team conducted additional data cleaning tasks and submitted a discrepancy log to RAND SRG, which was used

to continue data cleaning of the raw data files or to generate a query to market managers if SRG could not resolve the issue by reviewing the data. All staff followed a strict data safeguarding protocol to protect exposure of personal information.

Additional Data Collection Activities

In addition to collecting and analyzing quantitative assessment data during the Beta test, we collected qualitative feedback from Facility/Agency Staff and Research Nurses to augment our understanding of work flows, barriers, and challenges to implementing the assessment in practice, and the burden on staff. We administered surveys to Facility/Agency Staff, and conducted separate focus groups with both field staff and Research Nurses to collect this information. The approach taken for these activities is described below, and results of these activities are included in subsequent volumes of this report.

Facility/Agency Staff Survey

The purpose of the survey was to obtain information on several key stakeholder concerns identified through a review of comments received in response to the LTCH, SNF, IRF, and HH Notice of Proposed Rule Making; comments submitted as part of the subregulatory public comment period held in spring 2017; and other questions and comments received by CMS (e.g., via the PAC QI inbox, during Special Open Door Forums). Specifically, the survey was designed to collect information on perspectives of the clinical utility of Beta assessment candidate SPADEs, as well as the burden associated with collecting this information, for both the patient and the assessor. The survey also sought to understand factors associated with burden of data collection. Questions were included to ascertain the use of electronic medical records at the site, as well as the type and length of clinical experience of the assessor. The full survey instrument can be found in Appendix D.

The intended recipients of the survey were the Facility/Agency Staff assessors (approximate $n = 280$) who collected data at the Beta test sites ($n = 142$). The survey was administered through a web-based platform. Assessors were invited to participate via email, with up to three reminder emails sent as needed.

Descriptive statistics were calculated for the survey sample. For each candidate SPADE, we computed the mean of the reported scores for questions on clinical utility and burden. The average scores were plotted with bar charts. For SPADEs that included multiple items, each item was ranked relative to other items within the SPADE. Ranked items were displayed with stacked bar charts.

Facility/Agency Staff Focus Groups

Facility/Agency Staff who participated in the Beta test hold unique perspectives on the SPADEs, given their experiences with the assessment and data collection processes in the

different PAC settings. RAND conducted focus groups with field staff who participated in the Beta test. Capturing more detailed feedback from Facility/Agency Staff in their own words complements survey responses by providing important context for survey findings, illuminating any benefits or concerns that were not captured in the survey, and creating the opportunity for RAND to build rapport with providers on behalf of the SPADE initiative. Our goals were to compare the processes by which existing standardized assessments and the candidate SPADEs were completed, how electronic health records integrated with these processes, which Beta items were particularly easy or difficult to complete, which items were most clinically relevant and useful across PAC settings, and how the Facility/Agency Staff would like to see these items used in the future, if at all.

RAND sampled Facility/Agency Staff by location, selecting markets for participation with guidance of the Research Nurses, based on feasibility of data collection by focus group in that market. Markets in different regions with high engagement were prioritized. RAND conducted 90-minute, in-person focus groups in four markets: Boston and Durham in mid-June and Phoenix and St. Louis in mid-July. Each had five to eight participants that represented a diversity of PAC settings (SNFs, HHAs, IRFs, and LTCHs). RAND conducted one additional focus group with field staff electronically (via Adobe Connect) in Nashville because the Facility/Agency Staff were not able to meet in person.

Three RAND researchers alternately acted as moderators and note takers (with two attending each focus group). Focus groups followed a semistructured discussion guide that explored findings to date in greater depth. For example, some preliminary findings from the Beta test, like time to complete particular candidate SPADEs, warranted follow-up. Similarly, probes were generated to get additional information about select results of the Beta assessor survey. Focus groups were audio recorded and transcribed verbatim to facilitate directed content analysis.

Research Nurse Focus Groups

Given their experience coaching SNF, HHA, IRF, and LTCH staff, Research Nurses have unique perspectives on data collection during the Beta test as well as working knowledge of the existing data collection instruments across PAC settings. RAND conducted focus groups with Research Nurses to gather their perspectives on these topics, as well as to potentially inform future clinical testing by CMS. Specifically, RAND sought to gather Research Nurses' understanding of whether the concepts measured by Beta items were already collected and/or clinically useful and whether the candidate SPADEs adequately captured these concepts with minimal disruption to work flow in different PAC settings.

RAND conducted five electronic focus groups using Adobe Connect for the 28 Research Nurses. Each call included four to six Research Nurses from three markets, which were grouped by region. All electronic focus groups were conducted during the week of July 23, 2018. Research Nurses were typically invited to participate at times that overlapped with their regular calls with RAND market managers. Focus groups were 90 minutes in length, with discussions following a topic guide that incorporated findings of the facility/agency assessor survey as well

as other items of interest that were appropriate to explore in a focus group format. The same three RAND researchers that had conducted field staff focus groups again alternately acted as moderators and note takers (with two attending each focus group) and conducted analyses and summaries of field notes.

Chapter 7. Analytic Plan

Goals and Approach

For all communicative and noncommunicative admission assessments, there were three primary goals for the data analysis: (1) to compute descriptive statistics for candidate SPADEs, (2) to determine the feasibility of administration, and (3) to evaluate IRR between pairs of Research Nurses and Facility/Agency Staff assessors. Goals 1 and 2 applied also to communicative discharge assessments.

For some SPADEs, due to their complexity or the assessment design, additional analytic goals applied to the assessment data: (4) to understand the best look-back time frame and assessment days, (5) to identify the most appropriate versions of candidate SPADEs for cross-setting standardization, (6) to understand the psychometric properties of scalable candidate SPADEs such as from the PROMIS item banks, and (7) to determine the content validity of candidate SPADEs and distribution of scores.

Many of the analyses conducted during Beta testing focused on candidate SPADE characteristics and performance overall and by PAC setting. However, some analyses examined differences by specific groups of interest, such as by patient/resident and provider characteristics, which required the use of supplemental data. Specifically, in addition to the data collected by Research Nurses and Facility/Agency Staff, RAND requested the routine (or legacy) assessment data (OASIS, MDS, IRF-PAI, LCDS) from CMS that was collected concurrently by the PAC facilities/agencies and submitted to CMS to fulfill PAC prospective payment system requirements.

For patient/resident characteristics, we considered, when appropriate, differences by demographic groups (e.g., age, sex), clinical characteristics (e.g., total hip/knee joint replacement, septicemia or severe sepsis, heart failure and shock, stroke), length of stay, and prior setting. While the primary goal was to focus on patient characteristics that are comparably defined operationally across settings, analyses sometimes focused on setting-specific variables (e.g., for clinical characteristics that are not found in all settings, such as ventilator use) that could provide informative within-setting results. For provider characteristics, we considered, when appropriate, differences by whether a facility/agency is located in a rural or urban area, whether it is a freestanding facility or unit of a larger hospital (IRF/SNF), type of control (for-profit, nonprofit, governmental), region of the country (West, Midwest, South, Northeast), and site size (number of discharges per year in the form of groups based on within-setting percentiles).

Descriptive Statistics

Using data collected from Facility/Agency Staff and unpaired Research Nurse admission assessments, response frequencies were computed for each candidate SPADE to provide descriptive information pertaining to category response distributions overall, as well as by PAC setting. Additionally, basic descriptive statistics were computed for patient/resident characteristics (e.g., demographics, clinical subgroups), provider characteristics (e.g., site size, type of control, and geographic location), and admission and discharge assessments.

Where relevant, statistical comparisons were reported on tabulated data for the aforementioned groupings. Chi-squared tests were used to test whether distributions significantly varied by groups of interest. For continuous variables, differences in the distributions of scores for groups were tested using analysis of variance.

Feasibility

Missing Data

Missing data frequencies were computed overall, as well as by PAC setting, using data collected from Facility/Agency Staff and unpaired Research Nurse assessments in two stages. As there were situations that resulted in a module being skipped entirely (e.g., assessment window closed before interview portion could be completed), prior to computing missing data at the item level it was first confirmed that the module was attempted. The number of completed admission assessments at the module level is reported separately. In the second phase, missing data frequencies were examined for the number of cases in which the item had missing data or a category response was endorsed, indicating an inability to assess or no answer (e.g., “Unable to assess/no response”).

Tabulated missing data frequencies were analyzed using chi-squared tests to determine whether missing data distributions significantly varied by setting. For data with cell frequency counts greater than five, traditional chi-squared tests of independence were used. However, for instances with missing data frequencies less than five, alternative methods were implemented, such as Fisher’s exact test. Statistical comparisons were reported by PAC setting for admission and discharge.

Completion Time

Completion times were estimated based on the subset of assessments completed by Facility/Agency Staff, as they are the primary group of interest for evaluating the ease of administration. Time spent to complete the items was self-reported by Facility/Agency Staff as well as collected via time stamps recorded on tablets during the assessment administration. To maximize available data, these time sources were both considered. Records with missing self-report values and out-of-range tablet values (e.g., negative values) were excluded from time estimates. For cases where both self-reported and tablet time were available, the self-report

estimate was preferred unless it was clearly out of range (e.g., 140 minutes), in which case tablet time was used. The final data set was inspected for outliers to be excluded. Using an established method, values that fell more than 1.5 times the interquartile range above the third quartile were excluded from time estimate calculations.¹ Time to complete each candidate SPADE was evaluated overall and by PAC setting, patient/resident characteristics, provider characteristics, and admission and discharge. Completion times were compared across these groupings and statistically evaluated using analysis of variance methods. Additionally, to account for practice effects (e.g., medication reconciliation assessment being performed more quickly with practice), the analytic team performed sensitivity analyses wherein the first two assessments per Facility/Agency Staff were excluded. Distributions and mean time change with and without the exclusion of the first two assessments were examined.

Interrater Reliability

To determine whether items could be completed with acceptable IRR, we calculated the level of agreement between pairs of Research Nurses and Facility/Agency Staff assessors. For dichotomous data, we computed Cohen's kappa and overall agreement percentage. For ordinal item data, we computed weighted kappa, with Fleiss-Cohen (quadratic) weights. For continuous/approximately continuous data (more than five ordered categories), we computed interclass correlation coefficients (ICC). When ICCs were lower than expected (< 0.70), we also computed Pearson correlations and t-tests for mean differences to examine raw associations between rater scores while investigating problematic mean rater difference. Moreover, in order for kappa to be computed, the same response categories (e.g., 0, 1, 2) must be endorsed by both sets of raters. When this does not occur, kappa cannot be calculated. To address this issue, we used an established methodology wherein pseudocases for the missing response categories are created to produce a square table that allows for kappa to be computed. These pseudocases are given infinitesimally small weights so as to have no effect on the kappa coefficient.²

Additionally, for all data elements we report raw percent agreement as a supplemental/alternative index of IRR. It is important to note that kappa is sensitive to base/prevalence rates (e.g., high concentration of yeses or nos). When prevalence rates are extremely high or low, kappa is unstable, thus reducing confidence in obtained estimates.³ As such, the analytic team

¹ S. Seo, "A Review and Comparison of Methods for Detecting Outliers in Univariate Data Sets," thesis, University of Pittsburgh, 2006; J. W. Tukey, *Exploratory Data Analysis*, Reading, Mass.: Addison-Wesley Publishing Company, 1977.

² C. R. Stein, R. B. Devore, and B. E. Wojcik, "Calculation of the Kappa Statistic for Inter-Rater Reliability: The Case Where Raters Can Select Multiple Responses from a Large Number of Categories," *SUGI 30 Proceedings*, Philadelphia, Penn.: SAS Institute, Inc., April 2005; J. Uebersax, "Calculating Kappa with SAS," 2002.

³ D. V. Cicchetti, and A. R. Feinstein, "High Agreement but Low Kappa: II. Resolving the Paradoxes." *Journal of Clinical Epidemiology*, Vol. 43, No. 6, 1990, pp. 551–558; S. Xu, and M. F. Lorber, "Interrater Agreement Statistics with Skewed Data: Evaluation of Alternatives to Cohen's Kappa," *Journal of Consulting and Clinical*

compared observed prevalence rates for all data elements to prevalence rate ranges determined to yield adequate power. When prevalence rates fell within acceptable ranges, kappa coefficients were stable and interpretable. However, when rates fell outside acceptable bounds, we emphasized raw percent agreement as a measure of IRR.

Evaluative labels for different IRR ranges are conventionally used to interpret IRR results (slight/poor agreement = 0 to 0.2, fair = 0.21 to 0.40, moderate = 0.41 to 0.60, substantial/good = 0.61 to 0.80, excellent/almost perfect = 0.81 to 1.0). However, IRR values depend on a number of factors that may alter this interpretation. Most importantly, they are based on: (1) sample size of respondents; (2) the base rate or prevalence of a symptom/item response (or between-subjects variability); and (3) the magnitude of difference between the obtained IRR value and the “critical” value deemed unreliable. To remain conservative in our interpretation, we computed and reported the 95-percent confidence intervals (CIs) for each IRR estimate. If the lower bound of the CI fell below the moderate range (≤ 0.4), it may indicate unacceptably low IRR of the item(s) in the particular subgroup (e.g., IRF patients who experience pain) being assessed. In that case, no firm conclusions were drawn about the IRR for the particular subgroup assessed. We distinguished such results from “low reliability” coefficients to indicate a lack of evidence in favor of or against the reliability of the items.

IRR coefficients were computed overall and by each PAC setting type. Because IRR results are highly dependent on the prevalence of a given symptom/item response in the population being assessed (a.k.a. between-subjects variance), IRR coefficients were calculated for subpopulations but not directly compared across those subpopulations (i.e., PAC settings) where this prevalence is likely to differ. Thus, decisions about unacceptably low IRR values were made in reference to an external threshold (i.e., is 0.4 within the 95-percent CI?), and in the context of examining/reporting the additional factors of prevalence, between-subjects variance, and raw rater agreement.

Chart Review Look-Back Evaluation

For chart review look-back evaluation (admission and discharge) candidate SPADEs, goal 4 was to understand which look-back time frame (i.e., admission day one, admission day three, admission day five, admission day seven, and discharge day one, discharge day three) would be most appropriate for these items. That is, we sought to determine whether there are important differences in rates as a function of the look-back period, and whether this varies by setting. Note that below we use the term “occurrence” to refer to appliance presence (e.g., indwelling catheter) or service received (e.g., respiratory therapy). For admission chart reviews, the first day an occurrence is documented reflected the data point of interest and was coded as such for analytic

Psychology, Vol. 82, No. 6, 2014, pp. 1219–1227; T. Byrt, J. Bishop, and J. B. Carlin, “Bias, Prevalence and Kappa,” *Journal of Clinical Epidemiology*, Vol. 46, No. 5, 1993, pp. 423–429; M. L. McHugh, “Interrater Reliability: The Kappa Statistic,” *Biochemia Medica*, Vol. 22, No. 3, 2002, pp. 276–282.

purposes. Frequencies across days (one, three, five, seven) were converted to cumulative percentages to observe rates of increase. For admission data, cumulative percentages were used to compute the ratio of percent occurrence for each day (one, three, five) to the total percent occurrence by day seven. Like admission data, discharge data frequency distributions were converted to percentages and then to cumulative percentages. Cumulative percentages were used to compute the ratio of percent occurrence on discharge day three to the total percent occurrence on day one. The resulting values were used to determine whether rates of occurrence depend on the look-back period, and, if so, to identify the earliest chart review day that captures the majority of occurrences. Analyses were done overall and by PAC setting type.

Each candidate SPADE offers an ideal look-back option where the greatest information is captured within a shortest look-back period. While this is informative at the item level (e.g., respiratory treatment or indwelling catheter), we were primarily interested in an overall best look-back option for all candidate SPADEs. To that end, we examined the frequency distribution of ideal look-backs provided across candidate SPADE and identified the most frequent look-back supported by most items. Pros and cons for each look-back period are discussed, as well as how they vary by setting and patient/resident characteristics.

Repeat Assessment Evaluation

For repeat assessments, goal 4 was to identify which assessment days (i.e., day three, day five, day seven) are most appropriate for these items. Specifically, the goal was to understand whether it matters which day patients/residents are assessed. Thus, analyses examined whether there are significant and meaningful changes in rates or scores depending on the day a patient/resident is assessed, and whether that varies by setting. Results allow us to justify, empirically, whether it matters on which day the candidate SPADEs are administered, and offer a recommendation for an ideal assessment day for proposed SPADEs. Analyses were conducted overall and by setting. Repeat assessment candidate SPADEs included response frequencies and missing data information per assessment day.

In order to statistically compare observed differences on continuous data elements (or ordinal items with five or more ordered categories) over repeat assessments, within-subjects analysis of variance followed by pairwise comparisons were tested. Confidence intervals were also computed to supplement mean difference interpretations. Within and across PAC settings, this identified if and when, over the repeat assessments, rates/scores changed and whether this change was statistically significant.

For dichotomous or ordinal items with fewer than five response options, we include summary frequency tables for the number of cases with no change in responses across repeat assessment days as well as combinations of varying responses. These results are discussed to inform on overall rates of change. We also used chi-squared tests to examine whether there was an association between assessment day and response frequencies. Results statistically supported a conclusion regarding whether there were changes in responses depending on the assessment day.

For each candidate SPADE, we discuss the pros and cons for each assessment day and how pros and cons vary by PAC setting. Additionally, we provide a summary discussion of the pros and cons for each assessment day across all candidate SPADEs included in the repeat assessment design to determine if recommendations can be made regarding an overall ideal assessment day.

For Behavioral Signs and Symptoms, given the change in response options over repeat assessment days, we examined whether the day of assessment matters in terms of capturing the number of patients/residents exhibiting a behavioral symptom. In cases where there appears to be an effect, the extent of change was determined as the assessment time frame extends. These comparisons were done overall as well as by PAC setting to determine whether observed overall differences varied by setting.

Alternate Forms

For alternate forms, goal 5 was to identify meaningful differences across forms and to determine which form (version) of the candidate SPADE may be most appropriate for cross-setting standardization. Specifically, for PROMIS and Pain interview items, the goal was to determine whether different information is obtained depending on the period of time the patient/resident is asked to reflect on. For Expression and Understanding, current versions tested in Beta (two-item LCDS/IRF-PAI version and the three-item MDS version) assess the patient's/resident's ability to express him-/herself and understand others. However, a key distinction is that the two-item version conflates expression with speech clarity, but the three-item version separates these. Therefore, the goal was to evaluate and compare data collected on similar items from each version, as well as to assess the added utility of the speech clarity item.

We examined the distribution of scores on these items and overall scale scores (where relevant) as well as IRR to make comparisons between alternate forms. We examined both the statistical significance of mean differences between items on the different forms (when appropriate) and an effect size criterion to indicate meaningful differences, namely Cohen's $d = 0.2$.

In order to make comparisons between alternate forms, we also used legacy assessment data to look at within-patient/-resident comparisons by setting. For instance, with Pain items, we compared data collected from SNF residents who completed both the five-day time frame MDS version and the three-day time frame Beta items. For Expression and Understanding items, we compared data from IRF and LTCH patients who completed the three-item Beta version to their two-item IRF-PAI and LCDS versions. For SNFs, we compared data collected on residents who completed both the two-item Beta version and the three-item MDS version. We also used OASIS data to map both versions to OASIS Expression and Understanding. Additionally, we examined data collected from SNF residents who completed both the three-day Beta Behavioral Signs and Symptoms items and MDS seven-day items. For all these analyses, we used chi-squared tests for dichotomous/ordinal items and paired samples t-tests for continuous items.

Psychometric Properties

Goal 6 was to understand the psychometric properties of scalable data elements (e.g., PROMIS Anxiety). As an initial step, we examined basic item-level and scale reliability statistics. This included percent of responses at the floor and ceiling, skew, kurtosis, average inter-item correlations, item-total correlations, and Cronbach's alpha. These statistics are used to identify poorly functioning items that would preclude the application of item response theory (IRT) modeling. Additionally, in order to apply IRT, the assumption of unidimensionality must be met. As such, we computed and evaluated item-total correlations and explained common variance (ECV). Low values of item-total correlations (< 0.40) may indicate problematic items. For ECV, values ≥ 0.60 may indicate that an item set is sufficiently unidimensional. Items that form a sufficiently unidimensional scale were fit to the graded response model (GRM). The GRM can be used to estimate the discrimination and location parameter(s) of the items that are discussed.

Differential item functioning (DIF) was also examined. DIF occurs when items do not assess the same construct equally well, or vary in meaning, between different subgroups (e.g., gender) of patients. DIF can be identified when IRT-calibrated item parameters vary meaningfully across subgroups. We examined both nonuniform DIF (slopes differ between groups) and uniform DIF (intercepts differ between groups) using an ordinal logistic regression approach commonly used in the development of PROMIS measures. Key subgroups include gender, age (\leq or > 65), and PAC setting. Six comparisons between different regression models were used to determine whether there was (1) any DIF at all, (2) uniform DIF only, or (3) nonuniform DIF. We used the chi-squared likelihood-ratio statistic as the DIF detection criterion ($\alpha < 0.01$) and the pseudo- R^2 measure of magnitude (≤ 0.02) in model comparisons. This 0.02 value for R^2 magnitude is conventionally used to identify nontrivial DIF in the development of PROMIS instruments. Once DIF was identified, we graphically plotted two test characteristic curves (TCC), one with DIF and one assuming no DIF. We considered DIF notable if it had more than a small effect size (0.20) on the score distribution. Implications from IRT and DIF analyses are discussed in the context of possible item deletion due to poorly functioning items and/or the need to reduce time burden.

Validity

The focus of goal 7 was on the convergent validity of candidate SPADEs. That is, we focused on examining associations between Beta items that should be theoretically related based on literature and clinical practice. This was examined for each candidate SPADE, but ultimately analyses were prioritized for which the literature and/or clinical expertise provides strong support. We examined the associations within each PAC setting, given that the patient/resident populations are expected to differ across PAC settings. We examined other types of validity

where warranted, such as content validity (the extent to which a data element represents all facets of a given construct consistently across a variety of relevant populations).

Moreover, for properly functioning and thus valid data elements, we anticipated distributional differences on particular candidate SPADEs as a function of presence/absence of select clinical diagnoses. Goal 7 focused on evaluating and comparing frequency and score distributions by these identified groups. Merging Beta and legacy assessment data allowed us to identify conditions (e.g., total hip/knee replacement, septicemia or severe sepsis, heart failure and shock, stroke) that are common across and within PAC settings, and for which it makes sense to conduct tests of validity. Such analyses provided support for the validity of candidate SPADEs that differ across such characteristics in expected ways.

The type of statistical test conducted to assess validity depended on the comparability of the Beta and legacy instrument item. For ordinal items, we examined the strength of association using the Pearson correlation coefficient and tested the null hypothesis of no association versus the alternative of a linear association using a Mantel-Haenszel chi-squared test. Tests of association of nominal (unordered categorical) items were conducted using a standard chi-squared test.

Appendix A. Past Performance, Current Use, and Input Opportunities for Candidate SPADEs Tested in Beta

Table A.1. Past Performance, Current Use, and Input Opportunities for Candidate SPADEs Tested in Beta

Data Element (Information Source)	Past Performance	Current Use in PAC Settings	Input Opportunities
Cognitive Function			
BIMS (patient/resident interview)	The BIMS was tested in the PAC-PRD, where it was found to have substantial to almost perfect agreement for interrater reliability (kappa range of 0.71 to 0.91) when tested in all four PAC settings. ^a The BIMS has also been found to have excellent reliability and high correlation with the well-validated Modified Mini-Mental State (3MS) test in nursing home populations.	The BIMS is currently collected in the MDS 3.0 and the IRF-PAI.	TEP 1, PC 1, proposed rulemaking
CAM (multiple information sources)	The Short CAM has been shown to be effective in identifying delirium in validated research studies. ^b The four items selected for the Short CAM were found to best distinguish delirium from other types of cognitive impairment. In the MDS 3.0 national testing in nursing homes, the CAM had almost perfect interrater reliability agreement (kappa of 0.89 and 0.85). ^c When tested in the PAC-PRD, the CAM had substantial interrater reliability agreement for the “Inattention and Disorganized Thinking” questions (kappa range of 0.70 to 0.73), and the “Altered Level of Consciousness” question showed moderate agreement (kappa of 0.58). ^d	MDS 3.0 and LCDS currently use versions of the four-item CAM (Short CAM), but response options and wording differ slightly.	TEP 1, PC 1, proposed rulemaking
Staff Assessment of Mental Status (observation)	Studies testing in nursing home patients have shown it to have moderate interrater reliability ($r = 0.80$) ^e 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 ($r = 0.59$, $p < 0.05$). ^f	Included in the MDS 3.0.	TEP 2, PC 2, Alpha 2

Data Element (Information Source)	Past Performance	Current Use in PAC Settings	Input Opportunities
Cognitive Function			
Expression and Understanding (multiple information sources)	<p>Similar data elements were tested in the PAC-PRD and formed a composite <i>Communication</i> variable, which had substantial agreement for interrater reliability (kappa range of 0.74 to 0.80).^g</p> <p>In the national MDS 3.0 testing, Speech Clarity, Makes Self Understood, and Ability to Understand Others had almost perfect interrater reliability agreement (kappa range of 0.82 to 0.91).^h</p>	<p>Speech Clarity, Makes Self Understood, and Ability to Understand Others are currently included in the MDS 3.0.</p> <p>Expression of Ideas and Wants and Understanding Verbal Content are currently used in IRF-PAI and LCDS. OASIS-C2 uses a similar item, Speech and Oral (Verbal) Expression of Language, which differs in response option phrasing but uses similar tiers of difficulty.</p>	PC 1
Behavioral Signs and Symptoms: Presence and Frequency; Impact on Patient/Resident; Impact on Others; Rejection of Care (multiple information sources)	The data elements were tested in the national MDS 3.0 test and had almost perfect reliability (kappa of 0.90 and 0.94). Similar items were tested in PAC-PRD, but interrater reliability was not reported because of low incidence of disturbances.	The Behavioral Signs and Symptoms data elements are currently in the MDS 3.0. OASIS-C2 includes a similar data element, which records the frequency of disruptive behaviors.	TEP 2, PC 2, Alpha 2, proposed rulemaking
Mental Status			
Patient/Resident Mood Interview (PHQ-2 to -9; patient/resident interview)	The PHQ-9 has been validated in older adults, home health, skilled nursing facilities, and rehabilitation populations. The PHQ-9 has demonstrated ability to identify clinically important depression in adults of all ages, to make accurate diagnosis of major depression, to track severity of depression over time, and to monitor patient response to therapy. The PHQ-9 has also been shown to be a reliable and valid screening tool for detecting major depressive disorder (MDD) in patients with complex medical issues, including stroke and traumatic brain injury. The PHQ-2 was tested in the PAC-PRD and found to be reliable in Beta testing across the four PAC settings (kappas ranged from .74 to .91). ⁱ	The PHQ-9 and PHQ-2 are currently implemented in the MDS 3.0 and OASIS, respectively.	TEP 1, PC 1, Alpha 1, proposed rulemaking (PHQ-2)
Staff Assessment of Patient/Resident Mood (PHQ-9 OV; observation)	In the national validation of the MDS 3.0, facility nurses were able to complete the Staff Assessment of Resident Mood for 92% of the residents who did not complete the Resident Mood Interview. ^j	Currently included in the MDS 3.0.	TEP 2, PC 2, Alpha 2

Data Element (Information Source)	Past Performance	Current Use in PAC Settings	Input Opportunities
PROMIS Anxiety (patient/resident interview)	<p>PROMIS item banks have been developed for a large number of topics using rigorous methodology.</p> <p>Eleven anxiety items from the PROMIS item bank were tested in Alpha 2 feasibility testing. IRR ranged from 0.8 to 1, with most ratings falling between 0.95 to 1, indicating almost perfect agreement between assessors.</p> <p>Assessors' comments suggested that the Anxiety items were straightforward to administer and highly consistent across assessors; however, there were concerns that the items may be burdensome and that patients/residents were not strictly following the 7-day look-back period.</p>	PROMIS Anxiety items are not included in the current PAC assessments.	TEP 2, PC 2, Alpha 2
PROMIS Depression (patient/resident interview)	Because it was developed relatively recently, PROMIS Depression has not been as extensively validated as the other data elements included in Beta testing; it has, however, been shown to be sensitive to change in diverse clinical samples.	PROMIS Depression items are not included in the current PAC assessments.	Follow-up TEP webinar
Medical Conditions: Pain			
Pain interview: presence, frequency, severity, effect on sleep, interference with therapy- and non-therapy-related activities, relief (patient/resident interview)	<p>The data elements related to effect of pain on sleep and daily activities have been shown to enhance the evaluation of pain.</p> <p>The data element cluster in the MDS has demonstrated excellent reliability, with kappas of 0.96 and 0.97.</p> <p>A slightly modified version of the MDS 3.0 pain interview was tested in the PAC-PRD, in which the data elements demonstrated good reliability in all four PAC settings.</p> <p>Six pain interview data elements were included in Alpha 1 feasibility testing: Pain Frequency, Pain Effect on Sleep, Pain Interference with Activities (Therapy Activities), Pain Interference with Activities (Other Activities), Pain Severity, and Pain Relief. Alpha 1 feasibility testing demonstrated excellent interrater reliability and brief time to administer, and assessors' comments generally reflected that the pain items were straightforward to administer and highly consistent across assessors.</p>	<p>Interview-based data elements assessing pain are included in the MDS 3.0 and focus on pain presence, severity, effect on sleep, and effect on activities.</p> <p>OASIS asks about standardized pain assessment being conducted, and indication of severe pain.</p> <p>The IRF-PAI and LCDS have assessments of pain severity.</p>	TEP 1, PC 1, Alpha 1, TEP 2, PC 2

Data Element (Information Source)	Past Performance	Current Use in PAC Settings	Input Opportunities
Staff assessment of pain or distress (observation)	<p>The MDS data element cluster has demonstrated excellent reliability, with kappas of 0.94 and 0.96.</p> <p>IRR in Alpha 2 feasibility testing ranged from 0.69 to 1, indicating substantial to almost perfect agreement for the items. Assessors' feedback reflected that the pain items were straightforward, but somewhat challenging to administer, due to the time required for observation and the need to consult multiple data sources.</p>	<p>Observational data elements to assess pain are included in the OASIS-C2 and the MDS 3.0.</p> <p>The OASIS-C2 item queries the frequency of pain interfering with a patient's/resident's activity or movement.</p> <p>The set of data elements included in the MDS 3.0 document indicators of pain or possible pain across four types of behaviors, including nonverbal sounds, vocal complaints of pain, facial expressions, and protective body movements or postures.</p>	TEP 1, TEP 2, PC 2, Alpha 2
Impairments			
Hearing, Vision (multiple information sources)	<p>Slightly different versions of ability to hear and ability to see were tested in the PAC-PRD and showed substantial reliability across PAC settings (unweighted kappa of 0.78 for the hearing element; unweighted kappa of 0.74 for the vision element).</p> <p>In national MDS 3.0 testing, these data elements showed almost perfect interrater reliability (kappa range of 0.86–0.94).</p>	<p>MDS 3.0 currently assesses the hearing and vision data elements that are included in Beta testing.</p> <p>OASIS-C2 also assesses hearing and vision abilities.</p>	TEP 1, PC 1, proposed rulemaking
Bladder and Bowel Continence patient/resident perceived problem (patient/resident interview)	<p>In Alpha 1 testing, the items had almost perfect interrater reliability with kappas of 0.90 or higher.</p>	<p>Data elements to assess patient/resident perceived problem or burden with bladder and bowel incontinent events are not currently collected in PAC assessments.</p>	Alpha 1, PC 2

Data Element (Information Source)	Past Performance	Current Use in PAC Settings	Input Opportunities
Bladder and Bowel Continence Appliance use, frequency of events, need for assistance (chart review)	<p>In the national MDS 3.0 testing, the bladder and bowel data element cluster had almost perfect interrater reliability (kappa of 0.95).</p> <p>Similar items were also tested in PAC-PRD with kappas ranging from 0.60 to 0.90.</p> <p>Alpha 1 results suggest that items performed sufficiently well across settings. IRR on bladder and bowel items was substantial to almost perfect for items that could be assessed (kappas ranged from 0.61 to 0.9). Some items did not have sufficiently complete information, and kappas could not be calculated.</p>	<p>Similar items to those tested in Beta testing are included in current PAC assessments: bladder and bowel appliance use (MDS 3.0), bladder frequency of incontinent events (IRF-PAI, LCDS, MDS), and bowel frequency of incontinent events (OASIS, IRF-PAI, LCDS, MDS).</p>	TEP 1, Alpha 1, PC 2
Special Services, Treatments, and Interventions (SSTIs)			
Nutritional Approaches: Parental/IV Feeding, Feeding Tube, Mechanically Altered Diet, Therapeutic Diet (chart review)	<p>In the national MDS 3.0 testing, the nutritional approaches items exhibited almost perfect interrater reliability (kappa range of 0.89 to 0.96).</p>	<p>The nutritional approaches items are currently collected in MDS 3.0.</p>	PC 1, TEP 2, proposed rulemaking
Services and Treatments: Chemotherapy, Radiation, Oxygen Therapy, Suctioning, Tracheostomy Care, Invasive Mechanical Ventilator, Noninvasive Mechanical Ventilator, IV Medications, Transfusions, Dialysis, IV Access (chart review)	<p>A subset of the SSTIs (and without the subitems) was tested in PAC-PRD. Although IRR was not calculated, the use supports the feasibility of assessing SSTIs across PAC settings.</p> <p>In national MDS 3.0 testing, a slightly different suite of SSTIs showed almost perfect interrater reliability (kappa of 0.84 and 0.90).</p>	<p>The SSTIs (except for IV access and the subitems) are currently included in MDS 3.0.</p>	PC 1, TEP 2, proposed rulemaking
Other			
Care Preferences: Decisionmaking preferences, designated health care agent (multiple information sources)	<p>In the national MDS 3.0 testing, this item had almost perfect agreement (kappa of 0.99).</p> <p>The three care preferences data elements were tested in Alpha 1 feasibility testing. Agreement was uniformly high for paired observations across the items and across settings in Alpha 1. Interrater reliability was 0.90 or higher.</p> <p>The Advance Directive data element was tested in the PAC-PRD and found to be feasible to implement with a kappa of 0.72.</p>	<p>The MDS assesses the patient's/resident's preferences for family/significant other involvement in care discussions.</p>	TEP 1, Alpha 1, TEP 2, PC 2

Data Element (Information Source)	Past Performance	Current Use in PAC Settings	Input Opportunities
Medication Reconciliation (multiple information sources)	<p>MR items tested in Alpha 1 had the lowest IRR of any data elements tested (kappa coefficients ranged from 0.06 to 0.83), and assessors experienced difficulties completing the assessment. Therefore, MR showed considerable difficulty with reliability and feasibility.</p> <p>During Alpha 2 testing, medication reconciliation items demonstrated challenges and limitations that require revisions and retesting. Overall, there was variability in reliability across items and settings. Assessors suggested that further refinement to the items' instructions would improve the assessment.</p> <p>Based on Alpha 2 feedback, the burden of the MR items was reduced by eliminating the 7-day look-back period for the items on which medications were being taken, reducing the number of drug classes to six, and revising wording.</p>	<p>A similar item noting the classes of medications a patient/resident is taking is included in the MDS 3.0.</p> <p>However, none of the other MR items tested in Beta are in the current PAC assessments.</p>	TEP 1, Alpha 1, TEP 2, PC 2, Alpha 2
PROMIS Global Health (patient/resident interview)	Feedback from cognitive interviews on the PROMIS data elements on Global Health suggests that participants were unsure whether to respond based on their experience in the current setting or prior to admission since the data element includes items that ask about experiences "in general," "in the past 7 days," and with no specified time frame.	PROMIS Global Health is not included in any current PAC assessments, but is used on several large national and state surveys.	PC 2

SOURCES:

^a B. Gage et al., *The Development and Testing of the Continuity Assessment Record and Evaluation (CARE) Item Set: Final Report on Reliability Testing*, Vol. 2, Research Triangle Park, N.C.: RTI International, 2012.

^b S. K. Inouye et al., "The CAM-S: Development and Validation of a New Scoring System for Delirium Severity in 2 Cohorts," *Annals of Internal Medicine*, Vol. 160, No. 8, 2014, pp. 526–533.

^c D. Saliba, and J. Buchanan, *Development and Validation of a Revised Nursing Home Assessment Tool: MDS 3.0*, Santa Monica, Calif.: RAND Corporation, Contract No. 500-00-0027/Task Order #2, April 2008.

^d Gage et al., 2012.

^e R. Casten et al., "Psychometric Characteristics of the Minimum Data Set I: Confirmatory Factor Analysis," *Journal of the American Geriatric Society*, Vol. 46, No. 6, 1998, pp. 726–735.

^f M. P. Lawton et al., "Psychometric Characteristics of the Minimum Data Set II: Validity," *Journal of the American Geriatric Society*, Vol. 46, No. 6, 1998, pp. 736–744.

^g B. Gage et al., *Post-Acute Care Payment Reform Demonstration: Final Report*. Research Triangle Park, N.C.: RTI International, 2012.

^h Saliba and Buchanan, 2008.

ⁱ J. Klapow et al., "Psychological Disorders and Distress in Older Primary Care Patients: A Comparison of Older and Younger Samples," *Psychosomatic Medicine*, Vol. 64, No. 4, 2002, pp. 635–643; K. Kroenke, R. L. Spitzer, and J. B. Williams, "The PHQ-9: Validity of a Brief Depression Severity Measure," *Journal of General Internal Medicine*, Vol. 16, No. 9, 2001, pp. 606–613; B. Löwe et al., "Measuring Depression Outcome with a Brief Self-Report Instrument: Sensitivity to Change of the Patient Health Questionnaire (PHQ-9)," *Journal of Affective Disorders*, Vol. 81, No. 1, 2004, pp. 61–66; B. Ruo et al., "Depressive Symptoms and Health-Related Quality of Life: The Heart and Soul Study," *Journal of the American Medical Association*, Vol. 290, No. 2, 2003, pp. 215–221; and J. W. Williams et al., "Identifying Depression in Primary Care: A Literature Synthesis of Case-Finding Instruments," *General Hospital Psychiatry*, Vol. 24, No. 4, 2002, pp. 225–237.

^j Saliba and Buchanan, 2008.

Appendix B. Beta Research Nurse Training Agenda

October 2nd to 6th, 2017

Monday, October 2nd	Room 1224
8:00 am	Sign-in and pick-up materials
8:15 am	Tablet Assignment
8:45 am	Welcome and Introductions Review of Training Agenda and Activities <i>Sangeeta Ahluwalia, PhD</i>
9:15 am	Field Staff Training Plan: Schedule and Roles <i>Jason Etchegaray, PhD</i>
9:30 am	Beta Test Data Collection Procedures & Materials <i>Sangeeta Ahluwalia, PhD</i>
10:30 am	Break
10:45 am	Tablet Orientation and Walk-through <i>Jessica Phillips, MS, and Programming Team</i>
Communicative Assessment: Patient/Resident Interview Data Elements	
11:45 am	Data Element Training: Hearing, Vision, Expression, and Understanding <i>Lisa Newton, MSN, RN-BC, NEA-BC</i>
12:15 pm	Meet and Greet Luncheon
1:45 pm	Data Element Training: Cognition – Brief Interview for Mental Status (BIMS) <i>Cindy Lippman, RN</i>
2:15 pm	Data Element Training: PROMIS Global Health <i>Maria Edelen, PhD</i>
3:00 pm	Break
3:15 pm	Data Element Training: Pain Interview <i>Lisa Newton, MSN, RN-BC, NEA-BC</i>
4:00 pm	Wrap-up and Adjourn

Tuesday, October 3rd

Room 1224

8:00 am	Sign-in
8:30 am	Recap of yesterday's activities and plan for the day
8:45 am	Data Element Training: Mood – PHQ-2 to 9 <i>Robin Richardson, RN</i>
9:30 am	Data Element Training: Mood – PROMIS Depression; Anxiety <i>Maria Edelen, PhD</i>
10:00 am	Data Element Training: Care Preferences Interview <i>Julia Portale, MBA, MPH, LMSW</i>
10:15 am	Data Element Training: Continence Interview <i>Vicki Shier, PhD</i>
10:45 am	Break
Communicative Assessment: Multiple Information Sources Data Elements	
11:00 am	Data Element Training: Cognition – Confusion Assessment Method (CAM) <i>Ann Spenard, DNP, RN-BC</i>
11:30 am	Data Element Training: Behavioral Signs and Symptoms <i>Ann Spenard, DNP, RN-BC</i>
12:15 pm	Lunch
1:15 pm	Data Element Training: Medication Reconciliation <i>Robin Richardson, RN</i>
2:00 pm	Practice Exercise: Medication Reconciliation
2:45 pm	Data Element Training: Care Preferences <i>Julia Portale, MBA, MPH, LMSW</i>
3:00 pm	Break
3:15 pm	Data Element Training: Continence <i>Vicki Shier, PhD</i>
4:15 pm	Data Element Training: Special Services, Treatments, & Interventions (SSTI) <i>Ann Spenard, DNP, RN-BC</i>
5:00 pm	Wrap-up and Adjourn

Wednesday, October 4th Room 1224

8:00 am	Sign-in
8:30 am	Recap of yesterday's activities and plan for the day
8:45 am	Medication Reconciliation "Lessons Learned" Panel <i>Facilitators: Alpha 1 & Alpha 2 Research Nurses</i>
9:30 am	Role Play Introduction & Schedule Review <i>Sarah Dalton, MA</i>
9:45 am	Transition to Breakout Rooms
10:00 am	Role-Play #1 (Communicative Interview) Team A: Room 1232 (Facilitators: Deb Saliba, Ann Spenard, Brenda Karkos) Team B: Room 3312 (Facilitators: Jessica Phillips, Julia Rollison, Lisa Newton) Team C: Room 4312 (Facilitators: Jason Etchegaray, Julia Portale, Cindy Lippman) Team D: Room 5312 (Facilitators: Sangeeta Ahluwalia, Barb Gage, Robin Richardson)
12:00 pm	Transition to Main Room
12:15 pm	Lunch
1:15 pm	Debrief: Role-Play # 1
2:00 pm	Transition to Breakout Rooms
2:15 pm	Role-Play #2 (Communicative Interview) Team A: Room 1232 Team B: Room 3312 Team C: Room 4312 Team D: Room 5312
4:15 pm	Transition to Breakout Rooms
4:30 pm	Debrief Role-Play #2
5:15 pm	Wrap-up and Adjourn

Thursday, October 5th

Room 1224

8:00 am	Sign-in
8:30 am	Recap of yesterday's activities and plan for the day
	Non-Communicative Assessment
8:45 am	Data Element Training: Cognition – Staff Assessment of Mental Status (SAMS) <i>Brenda Karkos, MSN/MBA, RN, CHPN</i>
9:15 am	Data Element Training: Pain <i>Brenda Karkos, MSN/MBA, RN, CHPN</i>
9:45 am	Data Element Training: Mood – PHQ-OV <i>Robin Richardson, RN</i>
10:15 am	Break
10:30 am	Non-communicative assessment modeling: Mood <i>Robin Richardson, RN, and Cindy Lippman, RN</i>
11:00 am	Review and Reinforcement: Core assessment processes <i>Sangeeta Ahluwalia, PhD, and Jessica Phillips, MS</i>
11:30 am	Field Data Collection "Lessons Learned" Panel <i>Facilitators: Alpha 1 & Alpha 2 Research Nurses</i>
12:15 pm	Lunch
1:15 pm	Field Staff Training Plan: Schedule, Assignments, & Logistics <i>Jason Etchegaray, PhD</i>
1:45 pm	Training and Presenter Tips <i>Julia Rollison, PhD, MPH, and Lisa Newton, MSN, RN-BC, NEA-BC</i>
2:30 pm	Meet in Market Training Groups
3:00 pm	Break
3:15 pm	March October Madness – PAC Championships
4:30 pm	Wrap-up and Adjourn
4:45 pm	<i>Principal Trainer Meeting</i>

Friday, October 6th

Room 1224

8:00 am	Sign-in
8:15 am	Recap of yesterday's activities and plan for the day
8:45 am	Transition to Breakout Rooms
9:00 am	Practice Presentation Session #1 (Hearing, Vision, Expression, and Understanding; PROMIS: Global Health) Green Team: Room 1224 Red Team: Room 3312 Orange Team: Room 4212 Blue Team: Room 4312 Pink Team: Room 5212 White Team: Room 5312
10:30 am	Break
10:45 am	Practice Presentation Session #2 (Continance Interview; Behavioral Signs & Symptoms) Green Team: Room 1224 Red Team: Room 3312 Orange Team: Room 4212 Blue Team: Room 4312 Pink Team: Room 5212 White Team: Room 5312
12:30 pm	Lunch
1:30 pm	Transition to Breakout Rooms
1:45 pm	Practice Presentation Session #3 (Continance; Special Services, Treatments, and Interventions) Green Team: Room 1224 Red Team: Room 3312 Orange Team: Room 4212 Blue Team: Room 4312 Pink Team: Room 5212 White Team: Room 5312
3:30 pm	Break
3:45 pm	Debrief, next steps and distribute final materials
4:15 pm	Adjourn

Appendix C. Example Beta Field Training Agenda (Boston)

Wednesday, October 18th, 2017

8:00 am	Sign-in, pick-up materials including tablet assignment
8:30 am	Welcome and Introductions Review of Training Agenda and Activities <i>Emily Butcher, Regional Training Lead, RAND Corporation</i> Data Collection Overview <i>Patrick Orr, Regional Training Lead, RAND Corporation</i>
9:30 am	Tablet Orientation & Modeling <i>Emily Butcher, Regional Training Lead, RAND Corporation</i>
10:30 am	<i>Break</i>
Communicative Assessment: Patient/Resident Interview Data Elements	
10:40 am	Hearing, Vision, Expression and Understanding <i>Lynda Sheehan, RN</i>
11:10 am	Cognition Brief Interview for Mental Status (BIMS) <i>Ann Spenard, DNP, RN-BC</i>
11:35 am	PROMIS Global Health <i>Lynn Averill, RN</i>
12:00 pm	Pain <i>Brenda Karkos, MSN/MBA, RN, CHPN</i>
12:30 pm	<i>Lunch</i>

1:15 pm	Mood PHQ2 to 9 PROMIS Depression PROMIS Anxiety <i>Ann Spenard, DNP, RN-BC</i>
2:15 pm	Care Preferences <i>Brenda Karkos, MSN/MBA, RN, CHPN</i>
2:30 pm	Continence <i>Lynda Sheehan, RN</i>
2:45 pm	Break
2:55 pm	Tablet Reinforcement <i>Emily Butcher, Regional Training Lead, RAND Corporation</i>
3:25 pm	Role play (in pairs) #1 <i>Patrick Orr, Regional Training Lead, RAND Corporation</i>
4:15 pm	Debrief Role play/ Q&A <i>Patrick Orr, Regional Training Lead, RAND Corporation</i>

Thursday, October 19

8:00 am **Overview of the Day**
Patrick Orr, Regional Training Lead, RAND Corporation

Communicative Assessment: Multiple Information Sources Data Elements

8:10 am **Confusion Assessment Method (CAM)**
Ann Spenard, DNP, RN-BC

8:40 am **Behavioral Signs and Symptoms**
Lynn Averill, RN

9:10 am **Medication Reconciliation**
Ann Spenard, DNP, RN-BC

Communicative Assessment: Chart Review Data Elements

10:10 am *Break*

10:25 am **Care Preferences**
Brenda Karkos, MSN/MBA, RN, CHPN

10:40 am **Continence**
Lynda Sheehan, RN

11:00 am **Special Services, Treatments, and Interventions**
Lynda Sheehan, RN

11:30 am **Tablet Syncing Example**
Emily Butcher, Regional Training Lead, RAND Corporation

11:45 am **Modeling Patient Recruitment/Project Information Sheet**
Ann Spenard, DNP, RN-BC
Brenda Karkos, MSN/MBA, RN, CHPN

12:15 pm *Lunch*

Non-Communicative Patient/Resident Data Elements

1:00 pm **Core Assessment Process**
Emily Butcher, Regional Training Lead, RAND Corporation

1:30 pm	Cognition Staff Assessment of Mental Status <i>Brenda Karkos, MSN/MBA, RN, CHPN</i>
2:00 pm	Pain <i>Brenda Karkos, MSN/MBA, RN, CHPN</i>
2:30 pm	Mood Staff Assessment of Patient/Resident Mood (PHQ-9-OV) <i>Ann Spenard, DNP, RN-BC</i>
3:00 pm	Break
3:10 pm	Role play and debrief (in pairs) # 2 <i>Patrick Orr, Regional Training Lead, RAND Corporation</i>
3:50 pm	Role play and debrief (in pairs) # 3 <i>Patrick Orr, Regional Training Lead, RAND Corporation</i>
4:15 pm	Q&A Next Steps <i>Emily Butcher, Regional Training Lead, RAND Corporation</i>

Appendix D. Beta Field Staff Web-Based Survey Questions

The survey developed by Atlas Research is reproduced on the following pages.



Page 1 - Background

You are being asked to complete this survey because of your work conducting patient assessments in the Beta field test. The purpose of this survey is to collect information about your perceptions of the Beta assessment items, including (1) clinical utility of the assessment items; (2) ease and burden of data collection for the assessor and patient; and (3) factors that affected your ability to collect data; as well as processes for collecting and maintaining electronic patient/resident data at your facility. Most of the questions will apply to all data elements, with some more focused questions on select data elements included.

This survey should take you no more than 15 to 20 minutes to complete. Please note that if you exit the survey before completion, not all of your responses will be captured. You may choose to skip questions that you are not comfortable answering (unless required with an asterisk). Please be as honest as you can be in your responses, as they will be invaluable in improving the assessment. Results from the survey will be aggregated and your responses will be kept confidential and anonymous.

1. Market:* [\[Drop-Down\]](#)
 - Boston
 - Chicago
 - Dallas
 - Durham
 - Fort Lauderdale
 - Harrisburg
 - Houston
 - Kansas City
 - Los Angeles
 - Nashville
 - Philadelphia
 - Phoenix
 - San Diego
 - St. Louis

2. In which post-acute care setting are you currently administering the assessment?* [\[Drop-Down\]](#)
 - Home Health Agency (HHA)
 - Inpatient Rehabilitation Facility (IRF)
 - Long Term Care Hospital (LTCH)
 - Skilled Nursing Facility (SNF)

3. In which post-acute care setting(s) do you have experience?* *(Select all that apply.)*
 - Home Health Agencies (HHAs)
 - Inpatient Rehabilitation Facilities (IRFs)
 - Long Term Care Hospitals (LTCHs)
 - Skilled Nursing Facilities (SNFs)



4. How many years of experience do you have working in post-acute care?* [\[Drop-Down\]](#)
- <1 year
 - 1-2 years
 - 3-5 years
 - 6-10 years
 - >10 years
5. Please select your level of certification:*
- Licensed Practical Nurse (LPN) / Licensed Vocational Nurse (LVN)
 - Registered Nurse (RN)
 - Advanced Practice Registered Nurse (APRN)
 - Physical Therapist (PT)
 - Occupational Therapist (OT)
 - Other (please specify):

Page 2 - Electronic Medical Records

This section focuses on processes for collecting and maintaining electronic data at your facility.

6. When assessors at your facility/agency collect patient/resident information for the required assessment for your setting (e.g., MDS, IRF-PAI, LCDS, OASIS), which of the following do they use?
- An electronic form embedded in EMR
 - An electronic form outside of EMR
 - A paper form
 - A combination of electronic and paper forms
7. Would it make sense clinically and ease documentation burden if the items we are testing were included in an EMR?
- Yes
 - No

Page 3 - Clinical Utility

This section focuses on your perceptions of how clinically useful each Beta assessment data element is for patients/residents in the post-acute care setting.

8. Thinking generally about the data elements within the following categories, how clinically useful are these sections of the assessment?



	Not at all useful	Slightly useful	Somewhat useful	Moderately useful	Extremely useful
a) Hearing and Vision	<input type="checkbox"/>				
b) Expression and Understanding	<input type="checkbox"/>				
c) BIMS	<input type="checkbox"/>				
d) CAM	<input type="checkbox"/>				
e) Behavioral Signs and Symptoms	<input type="checkbox"/>				
f) Pain Interview	<input type="checkbox"/>				
g) PHQ-2 to 9 Interview	<input type="checkbox"/>				
h) PROMIS Depression	<input type="checkbox"/>				
i) Anxiety	<input type="checkbox"/>				
j) Care Preferences	<input type="checkbox"/>				
k) Continence	<input type="checkbox"/>				
l) PROMIS Global Health	<input type="checkbox"/>				
m) Medication Reconciliation	<input type="checkbox"/>				
n) Nutritional Approaches	<input type="checkbox"/>				
o) Special Services, Treatments, and Interventions	<input type="checkbox"/>				
p) Staff Assessment of Cognitive Status (non-communicative)	<input type="checkbox"/>				
q) Staff Assessment of Pain (non-communicative)	<input type="checkbox"/>				
r) Staff Assessment of Mood (non-communicative)	<input type="checkbox"/>				

9. Please review the following list of items from the **Bladder Continence** data element and **rank them in terms of clinical utility**, with **1 being the item you feel is the most clinically useful** and **6 being the least clinically useful**.

Rank the items below, using numeric values starting with 1.

	Rank (1-6)
a) Perceived Problem or Burden with Bladder Incontinent Events	
b) Bladder Appliance Use	
c) Bladder Appliance Use – Current Setting	
d) Bladder Appliance Use – Primary Reason for Catheter Placement	



e) Bladder Appliance Use – Need for Assistance in Appliance Management	
f) Bladder – Frequency of Incontinent Events	

10. Please review the following list of items from the **Bowel Continence** data element and **rank them in terms of clinical utility**, with **1 being the item you feel is the most** clinically useful and **5 being the least** clinically useful.

Rank the items below, using numeric values starting with 1.

	Rank (1-5)
a) Perceived Problem or Burden with Bowel Incontinent Events	
b) Bowel Appliance Use	
c) Bowel Appliance Use – Current Setting	
d) Bowel Appliance Use – Need for Assistance in Appliance Management	
e) Bowel – Frequency of Incontinent Events	

11. Please review the following list of items from the **Pain Interview** data element and **rank them in terms of clinical utility**, with **1 being the item you feel is the most** clinically useful and **7 being the least** clinically useful.

Rank the items below, using numeric values starting with 1.

	Rank (1-7)
a) Pain Presence	
b) Pain Frequency	
c) Pain Effect on Sleep	
d) Pain Interference, Therapy Activities	
e) Pain Interference, Other Activities	
f) Pain Severity – Numeric Pain Scale	
g) Pain Relief	

12. Thinking about the following list of items from **Special Services, Treatments, and Interventions**, how clinically useful are each of these items?

	Not at all useful	Slightly useful	Somewhat useful	Moderately useful	Extremely useful
a) Chemotherapy	<input type="checkbox"/>				
b) Radiation	<input type="checkbox"/>				
c) Oxygen therapy	<input type="checkbox"/>				
d) Suctioning	<input type="checkbox"/>				
e) Tracheostomy care	<input type="checkbox"/>				
f) Invasive mechanical ventilator	<input type="checkbox"/>				
g) Non-invasive mechanical ventilator	<input type="checkbox"/>				



h) IV medications	<input type="checkbox"/>				
i) Transfusions	<input type="checkbox"/>				
j) Dialysis	<input type="checkbox"/>				
k) IV access	<input type="checkbox"/>				

13. [Open-Ended] Are there other data elements that were not included in the Beta assessment items that you believe would be clinically useful?

Page 4 - Ease, Burden, and Factors Affecting Information Collection

This section focuses on your perceptions of how difficult it was to collect information during the Beta assessment and how burdensome information collection was for patients/residents in your current post-acute care setting. It also asks about the factors that contributed to difficulty in collecting information.

14. Thinking generally about the data elements within the following categories, how difficult was it for you, as the assessor, to collect information for the following sections of the assessment?

	Not at all difficult	Slightly difficult	Somewhat difficult	Moderately difficult	Extremely difficult
a) Hearing and Vision	<input type="checkbox"/>				
b) Expression and Understanding	<input type="checkbox"/>				
c) BIMS	<input type="checkbox"/>				
d) CAM	<input type="checkbox"/>				
e) Behavioral Signs and Symptoms	<input type="checkbox"/>				
f) Pain Interview	<input type="checkbox"/>				
g) PHQ-2 to 9 Interview	<input type="checkbox"/>				
h) PROMIS Depression	<input type="checkbox"/>				
i) Anxiety	<input type="checkbox"/>				
j) Care Preferences	<input type="checkbox"/>				
k) Continence Interview	<input type="checkbox"/>				
l) Continence Chart Review	<input type="checkbox"/>				
m) PROMIS Global Health	<input type="checkbox"/>				
n) Medication Reconciliation	<input type="checkbox"/>				
o) Nutritional Approaches	<input type="checkbox"/>				
p) Special Services, Treatments, and Interventions	<input type="checkbox"/>				



q) Staff Assessment of Cognitive Status (non-communicative)	<input type="checkbox"/>				
r) Staff Assessment of Pain (non-communicative)	<input type="checkbox"/>				
s) Staff Assessment of Mood (non-communicative)	<input type="checkbox"/>				

15. For the data elements or categories of data elements that were **moderately** or **extremely difficult** to collect information for, what factors impacted your ability to collect information? *(Select all that apply.)*

- Availability of data within the information source (e.g., relevant information not available in chart)
- Could not complete the item as instructed (e.g., getting access to patient/resident, confusion, annoyance, comfort level)
- Timing constraints (e.g., limited assessor availability, limited patient/resident availability, data elements took longer than expected, timeliness of information sources)
- Environmental constraints (e.g., lack of physical space, lack of privacy, interruptions, background noise)
- Challenges interpreting patient response or available information
- Other (please describe):

16. We understand some items may be perceived as burdensome for patients and residents. Thinking generally about the data elements within the following categories, how burdensome would you say it was to the patient/resident to provide information for the following sections of the assessment?

	Not at all burdensome	Slightly burdensome	Somewhat burdensome	Moderately burdensome	Extremely burdensome
a) BIMS	<input type="checkbox"/>				
b) Pain Interview	<input type="checkbox"/>				
c) PHQ-2 to 9 interview	<input type="checkbox"/>				
d) PROMIS Depression	<input type="checkbox"/>				
e) Anxiety	<input type="checkbox"/>				
f) Care Preferences	<input type="checkbox"/>				
g) Continence interview	<input type="checkbox"/>				
h) PROMIS Global Health	<input type="checkbox"/>				



Page 5 - Ease, Burden, and Factors Affecting Information Collection (cont.)

17. Please review the following list of items from the **Bladder Continence** data element and **rank them in terms of burden to the assessor**, with **1 being the item you feel is most burdensome** and **6 being the least** burdensome.

Rank the items below, using numeric values starting with 1.

	Rank (1-6)
a) Perceived Problem or Burden with Bladder Incontinent Events	
b) Bladder Appliance Use	
c) Bladder Appliance Use – Current Setting	
d) Bladder Appliance Use – Primary Reason for Catheter Placement	
e) Bladder Appliance Use – Need for Assistance in Appliance Management	
f) Bladder – Frequency of Incontinent Events	

18. For the **Bladder Continence** items that were **most burdensome**, what factors impacted your ability to collect information? (*Select all that apply.*)

- Availability of data within the information source (e.g., relevant information not available in chart)
- Could not complete the item as instructed (e.g., getting access to patient/resident, confusion, annoyance, comfort level)
- Timing constraints (e.g., limited assessor availability, limited patient/resident availability, data elements took longer than expected, timeliness of information sources)
- Environmental constraints (e.g., lack of physical space, lack of privacy, interruptions, background noise)
- Challenges interpreting patient response or available information
- Other (please describe):

19. Please review the following list of items from the **Bowel Continence** data element and **rank them in terms of burden to the assessor**, with **1 being the item you feel is most burdensome** and **5 being the least** burdensome.

Rank the items below, using numeric values starting with 1.

	Rank (1-5)
a) Perceived Problem or Burden with Bowel Incontinent Events	
b) Bowel Appliance Use	
c) Bowel Appliance Use – Current Setting	
d) Bowel Appliance Use – Need for Assistance in Appliance Management	
e) Bowel – Frequency of Incontinent Events	



20. For the **Bowel Continence** items that were **most burdensome**, what factors impacted your ability to collect information? *(Select all that apply.)*

- Availability of data within the information source (e.g., relevant information not available in chart)
- Could not complete the item as instructed (e.g., getting access to patient/resident, confusion, annoyance, comfort level)
- Timing constraints (e.g., limited assessor availability, limited patient/resident availability, data elements took longer than expected, timeliness of information sources)
- Environmental constraints (e.g., lack of physical space, lack of privacy, interruptions, background noise)
- Challenges interpreting patient response or available information
- Other (please describe):

Page 6 - Ease, Burden, and Factors Affecting Information Collection (cont.)

21. Please review the following list of items from the **Pain** data elements and **rank them in terms of burden to the assessor**, with **1 being the item you feel is most burdensome** and **7 being the least burdensome**.

Rank the items below, using numeric values starting with 1.

	Rank (1-7)
a) Pain Presence	
b) Pain Frequency	
c) Pain Effect on Sleep	
d) Pain Interference, Therapy Activities	
e) Pain Interference, Other Activities	
f) Pain Severity – Numeric Pain Scale	
g) Pain Relief	

22. For the **Pain** items that were **most burdensome**, what factors impacted your ability to collect information? *(Select all that apply.)*

- Availability of data within the information source (e.g., relevant information not available in chart)
- Could not complete the item as instructed (e.g., getting access to patient/resident, confusion, annoyance, comfort level)
- Timing constraints (e.g., limited assessor availability, limited patient/resident availability, data elements took longer than expected, timeliness of information sources)
- Environmental constraints (e.g., lack of physical space, lack of privacy, interruptions, background noise)
- Challenges interpreting patient response or available information
- Other (please describe):



23. Thinking about the following list of items from **Special Services, Treatments, and Interventions** how burdensome was it for you, as the assessor, to collect information?

	Not at all burdensome	Slightly burdensome	Somewhat burdensome	Moderately burdensome	Extremely burdensome
a) Chemotherapy	<input type="checkbox"/>				
b) Radiation	<input type="checkbox"/>				
c) Oxygen therapy	<input type="checkbox"/>				
d) Suctioning	<input type="checkbox"/>				
e) Tracheostomy care	<input type="checkbox"/>				
f) Invasive mechanical ventilator	<input type="checkbox"/>				
g) Non-invasive mechanical ventilator	<input type="checkbox"/>				
h) IV medications	<input type="checkbox"/>				
i) Transfusions	<input type="checkbox"/>				
j) Dialysis	<input type="checkbox"/>				
k) IV access	<input type="checkbox"/>				

24. For the **Special Services, Treatments, and Intervention** items that were **moderately** or **extremely difficult** to collect information for, what factors impacted your ability to collect information? (*Select all that apply.*)

- Availability of data within the information source (e.g., relevant information not available in chart)
- Could not complete the item as instructed (e.g., getting access to patient/resident, confusion, annoyance, comfort level)
- Timing constraints (e.g., limited assessor availability, limited patient/resident availability, data elements took longer than expected, timeliness of information sources)
- Environmental constraints (e.g., lack of physical space, lack of privacy, interruptions, background noise)
- Challenges interpreting patient response or available information
- Other (please describe):

Page 7 – Additional Comments

25. [Open-Ended] Please describe any workarounds assessors are using to make data collection for this pilot assessment tool and/or for your facility’s current assessment tool easier (e.g., reordering data elements to make it easier to complete the assessment, collecting information on paper and then copying it over to the tablet):

26. [Open-Ended] Please provide any additional questions or comments related to the assessment tool (if any):



ATLAS
RESEARCH

Thank you for completing the survey. As a reminder, if you have questions about how to conduct the assessment or are running into problems in the field, please contact your research nurse.

References

- Alexopoulos, G. S., et al., "Pharmacotherapy of Depression in Older Patients: A Summary of the Expert Consensus Guidelines," *Journal of Psychiatric Practice*, Vol. 7, No. 6, 2001, pp. 361–376.
- Amtmann, D., et al., "Pain Affects Depression Through Anxiety, Fatigue, and Sleep in Multiple Sclerosis," *Rehabilitation Psychology*, Vol. 60, No. 1, 2015, p. 81.
- Areán, P. A., and B. L. Cook, "Psychotherapy and Combined Psychotherapy/Pharmacotherapy for Late Life Depression," *Biological Psychiatry*, Vol. 52, No. 3, 2002, pp. 293–303.
- Avenanti, A., et al., "Transcranial Magnetic Stimulation Highlights the Sensorimotor Side of Empathy for Pain," *Nature Neuroscience*, Vol. 8, No. 7, 2005, pp. 955–960.
- Ballard, C., et al., "Management of Agitation and Aggression Associated with Alzheimer's Disease: Controversies and Possible Solutions," *Current Opinion in Psychiatry*, Vol. 22, No. 6, 2009, pp. 532–540.
- Banerjee, S., and A. Macdonald, "Mental Disorder in an Elderly Home Care Population: Associations with Health and Social Service Use," *British Journal of Psychiatry*, Vol. 168, No. 6, 1996, pp. 750–756.
- Bao, Y., et al., "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.
- Barnsteiner, J. H., "Medication Reconciliation: Transfer of Medication Information Across Settings—Keeping It Free from Error," *Journal of Infusion Nursing*, Vol. 28, No. 2, Suppl., 2005, pp. 31–36.
- Bruce, M. L., G. J. McAvay, and P. J. Raue, "Major Depression in Elderly Home Health Care Patients," *American Journal of Psychiatry*, Vol. 159, No. 8, 2002, pp. 1367–1374.
- Byrt, T., J. Bishop, and J. B. Carlin, "Bias, Prevalence and Kappa," *Journal of Clinical Epidemiology*, Vol. 46, No. 5, 1993, pp. 423–429.
- Cadogan, M. P., et al., "A Standardized Quality Assessment System to Evaluate Pain Detection and Management in the Nursing Home," *Journal of Post-Acute and Long-Term Care Medicine*, Vol. 7, No. 3, Supplement, 2006, pp. S11–S19. As of February 6, 2019: [https://www.jamda.com/article/S1525-8610\(05\)00685-7/abstract](https://www.jamda.com/article/S1525-8610(05)00685-7/abstract).

Carson, A., et al., “The Outcome of Depressive Disorders in Neurology Patients: A Prospective Cohort Study,” *Journal of Neurology, Neurosurgery and Psychiatry*, Vol. 74, No. 7, 2003, pp. 893–896.

Casten, R., et al., “Psychometric Characteristics of the Minimum Data Set I: Confirmatory Factor Analysis,” *Journal of the American Geriatric Society*, Vol. 46, No. 6, 1998, pp. 726–735.

Centers for Medicare & Medicaid Services, *Development and Maintenance of Post-Acute Care Cross-Setting Standardized Patient Assessment Data: Technical Expert Panel Summary/Expert Input Report*, June 2016. As of February 6, 2019:
https://www.rand.org/content/dam/rand/pubs/working_papers/WR1100/WR1187/RAND_WR1187.pdf

———, *Development of Functional Outcome Quality Measures for Skilled Nursing Facilities (SNFs): Public Comment Summary Report*, September 2016. As of February 6, 2019:
<https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Development-of-Functional-Outcome-Quality-Measures-for-SNFs-Public-Comment-Summary-Report.pdf>

———, *Development and Maintenance of Post-Acute Care Cross-Setting Standardized Patient Assessment Data: Technical Expert Panel Summary Report (Second Convening)*, March 2017. As of February 6, 2019:
<https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/RAND-IMPACT-TEP-Second-Convening-Final-Report-March-2017.pdf>

———, “National Field Test Assessment Protocol: Admission,” November 2017. As of January 31, 2019:
https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/National-Field-Test-Assessment-Protocol_Admission.pdf

———, “National Field Test Assessment Protocol: Discharge,” November 2017. As of January 31, 2019:
https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/National-Field-Test-Assessment-Protocol_Discharge.pdf

_____, “National Field Test Assessment Protocol: Non-Communicative,” November 2017. As of January 31, 2019:

https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/National-Field-Test-Assessment-Protocol_Non-Communicative.pdf

_____, *Development and Maintenance of Post-Acute Care Cross-Setting Standardized Assessment Data: Public Comment Summary Report 2*, January 2018. As of February 6, 2019:

https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Public-Comment-Summary-Report_Standardized-Patient-Assessment-Data-Element-Work_PC2.pdf

Cicchetti, D. V., and A. R. Feinstein, “High Agreement but Low Kappa: II. Resolving the Paradoxes,” *Journal of Clinical Epidemiology*, Vol. 43, No. 6, 1990, pp. 551–558.

Diefenbach, G. J., D. F. Tolin, and C. M. Gilliam, “Impairments in Life Quality Among Clients in Geriatric Home Care: Associations with Depressive and Anxiety Symptoms,” *International Journal of Geriatric Psychiatry*, Vol. 27, No. 8, 2012, pp. 828–835.

DiMatteo, M. R., H. S. Lepper, and T. W. Croghan, “Depression Is a Risk Factor for Noncompliance with Medical Treatment: Meta-Analysis of the Effects of Anxiety and Depression on Patient Adherence,” *Archives of Internal Medicine*, Vol. 160, No. 14, 2000, pp. 2101–2107.

Dougherty, L. M., et al., “Effects of Exposure to Aggressive Behavior on Job Satisfaction of Health Care Staff,” *Journal of Applied Gerontology*, Vol. 11, No. 2, 1992, pp. 160–172.

Edelen, M. O., et al., *Development and Maintenance of Standardized Cross Setting Patient Assessment Data for Post-Acute Care: Summary Report of Findings from Alpha 1 Pilot Testing*, Santa Monica, Calif.: RAND Corporation, RR-1895-CMS, March 2017. As of February 6, 2019:

https://www.rand.org/pubs/research_reports/RR1895.html

Edelen, M. O., et al., *Development and Maintenance of Standardized Cross-Setting Patient Assessment Data for Post-Acute Care: Summary Report of Findings from Alpha 2 Pilot Testing*, Santa Monica, Calif.: RAND Corporation, EP-67619, June 2018. As of February 6, 2019:

https://www.rand.org/pubs/external_publications/EP67619.html

Ell, K., et al., “Routine PHQ-9 Depression Screening in Home Health Care: Depression Prevalence, Clinical and Treatment Characteristics, and Screening Implementation,” *Home Health Care Services Quarterly*, Vol. 24, No. 4, 2006, pp. 1–19.

- Engle, V. F., M. J. Graney, and A. Chan, "Accuracy and Bias of Licensed Practical Nurse and Nursing Assistant Ratings of Nursing Home Residents' Pain," *Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, Vol. 56, 2001, pp. M405–M411.
- Ferrell, B. A., "Assessing Pain in the Elderly," *The Consultant Pharmacist*, Vol. 25, 2010, Suppl. A: pp. 5–10.
- Freeman, E. E., et al., "Visual Field Loss Increases the Risk of Falls in Older Adults: The Salisbury Eye Evaluation," *Investigative Ophthalmology and Visual Science*, Vol. 48, No. 10, 2007, pp. 4445–4450.
- Gage, B., et al., *The Development and Testing of the Continuity Assessment Record and Evaluation (CARE) Item Set: Final Report on Reliability Testing*, Vol. 2, Research Triangle Park, N.C.: RTI International, 2012. As of February 6, 2019: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/The-Development-and-Testing-of-the-Continuity-Assessment-Record-and-Evaluation-CARE-Item-Set-Final-Report-on-Reliability-Testing-Volume-2-of-3.pdf>
- Gage, B., et al., *Post-Acute Care Payment Reform Demonstration: Final Report*, Research Triangle Park, NC: RTI International. 2012. As of February 6, 2019: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/Research-Reports-Items/PAC_Payment_Reform_Demo_Final.html
- Gleason, K. M., et al., "Reconciliation of Discrepancies in Medication Histories and Admission Orders of Newly Hospitalized Patients," *American Journal of Health-System Pharmacy*, Vol. 61, No. 16, 2004, pp. 1689–1695.
- Harris, Y., and J. K. Cooper, "Depressive Symptoms in Older People Predict Nursing Home Admission," *Journal of the American Geriatrics Society*, Vol. 54, No. 4, 2006, pp. 593–597.
- Hughes, R. G., et al., "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*, Vol. AHRQ, Pub. No. 08-0043, Rockville, Md.: Agency for Healthcare Research and Quality, 2008, As of February 6, 2019: <https://archive.ahrq.gov/professionals/clinicians-providers/resources/nursing/resources/nurseshdbk/nurseshdbk.pdf>
- Husebo, B. S., C. Ballard, J. Cohen-Mansfield, R. Seifert, and D. Aarsland, "The Response of Agitated Behavior to Pain Management in Persons with Dementia," *American Journal of Geriatric Psychiatry*, Vol. 22, No. 7, 2014, pp. 708–717.
- Husebo, B. S., C. Ballard, F. Fritze, R. K. Sandvik, and D. Aarsland, "Efficacy of Pain Treatment on Mood Syndrome in Patients with Dementia: A Randomized Clinical Trial," *International Journal of Geriatric Psychiatry*, Vol. 29, No. 8, 2014, pp. 828–836.

- Iliffe, S., et al., “The Relationship Between Pain Intensity and Severity and Depression in Older People: Exploratory Study,” *BMC Family Practice*, Vol. 10, No. 1, 2009, 54.
- Inouye, S. K., et al., “The CAM-S: Development and Validation of a New Scoring System for Delirium Severity in 2 Cohorts,” *Annals of Internal Medicine*, Vol. 160, No. 8, 2014, pp. 526–533.
- Institute of Medicine, *Preventing Medication Errors*, Washington, D.C.: The National Academies Press, 2007. As of February 6, 2019:
<https://www.nap.edu/read/11623/chapter/1>
- Ishii, S., J. E. Streim, and D. Saliba, “Potentially Reversible Resident Factors Associated with Rejection of Care Behaviors,” *Journal of the American Geriatrics Society*, Vol. 58, No. 9, 2010, pp. 1693–1700.
- , “A Conceptual Framework for Rejection of Care Behaviors: Review of Literature and Analysis of Role of Dementia Severity,” *Journal of the American Medical Directors Association*, Vol. 13, No. 1, 2012, pp. 11–23.
- Ivers, R. Q., et al., “Visual Impairment and Risk of Hip Fracture,” *American Journal of Epidemiology*, Vol. 152, No. 7, 2000, pp. 633–639.
- The Joint Commission, *Comprehensive Accreditation Manual for Hospitals*, Chicago, Ill.: The Joint Commission, 2015.
- Jones, R. N., E. R. Marcantonio, and T. Rabinowitz, “Prevalence and Correlates of Recognized Depression in U.S. Nursing Homes,” *Journal of the American Geriatrics Society*, Vol. 51, No. 10, 2003, pp. 1404–1409.
- Jordan, A., et al., “The Utility of PAINAD in Assessing Pain in a UK Population with Severe Dementia,” *International Journal of Geriatric Psychiatry*, Vol. 26, No. 2, 2011, pp. 118–126.
- Kane, K. D., B. P. Yochim, and P. A. Lichtenberg, “Depressive Symptoms and Cognitive Impairment Predict All-Cause Mortality in Long-Term Care Residents,” *Psychology and Aging*, Vol. 25, No. 2, 2010, 446.
- Katon, W., J. E. Russo, M. Von Korff, E. H. B. Lin, E. Ludman, and P.S. Ciechanowski, “Long-Term Effects on Medical Costs of Improving Depression Outcomes in Patients with Depression and Diabetes,” *Diabetes Care*, Vol. 31, No. 6, 2008, pp. 1155–1159.
- Katon, W., M. Von Korff, P., Ciechanowski, et al., “Behavioral and Clinical Factors Associated with Depression Among Individuals with Diabetes,” *Diabetes Care*, Vol. 27, No. 4, 2004, pp. 914–920.

- Klapow, J., et al., "Psychological Disorders and Distress in Older Primary Care Patients: A Comparison of Older and Younger Samples," *Psychosomatic Medicine*, Vol. 64, No. 4, 2002, pp. 635–643.
- Kroenke, K., et al., "The Association of Depression and Pain with Health-Related Quality of Life, Disability, and Health Care Use in Cancer Patients," *Journal of Pain and Symptom Management*, Vol. 40, No. 3, 2010, pp. 327–341.
- Kroenke, K., R. L. Spitzer, and J. B. Williams, "The PHQ-9: Validity of a Brief Depression Severity Measure," *Journal of General Internal Medicine*, Vol. 16, No. 9, 2001, pp. 606–613.
- Landefeld, C. S., et al., "National Institutes of Health State-of-the-Science Conference Statement: Prevention of Fecal and Urinary Incontinence in Adults," *Annals of Internal Medicine*, Vol. 148, No. 6, 2008, pp. 449–458.
- Lawton, M. P., et al., "Psychometric Characteristics of the Minimum Data Set II: Validity," *Journal of the American Geriatric Society*, Vol. 46, No. 6, 1998, pp. 736–744.
- Lebowitz, B. D., et al., "Diagnosis and Treatment of Depression in Late Life: Consensus Statement Update," *Journal of the American Medical Association*, Vol. 278, No. 14, 1997, pp. 1186–1190.
- Lequerica, A. H., et al., "Agitation in Acquired Brain Injury: Impact on Acute Rehabilitation Therapies," *Journal of Head Trauma Rehabilitation*, Vol. 22, No. 3, 2007, pp. 177–183.
- Lin, W.-C., et al., "Measuring Pain Presence and Intensity in Nursing Home Residents," *Journal of the American Medical Directors Association*, Vol. 7, No. 3, 2006, pp. 147–153.
- Löwe, B., et al., "Measuring Depression Outcome with a Brief Self-Report Instrument: Sensitivity to Change of the Patient Health Questionnaire (PHQ-9)," *Journal of Affective Disorders*, Vol. 81, No. 1, 2004, pp. 61–66.
- McHugh, J., A.-M. Casey, and B. Lawlor, "Psychosocial Correlates of Aspects of Sleep Quality in Community-Dwelling Irish Older Adults," *Aging and Mental Health*, Vol. 15, No. 6, 2011, pp. 749–755.
- McHugh, M. L., "Interrater Reliability: The Kappa Statistic," *Biochemia Medica*, Vol. 22, No. 3, 2002, pp. 276–282.
- Morley, J. E., "Weight Loss in the Nursing Home," *Journal of the American Medical Directors Association*, Vol. 8, No. 4, 2007, pp. 201–204.
- Moscou-Jackson, G., et al., "Acute Pain and Depressive Symptoms: Independent Predictors of Insomnia Symptoms Among Adults with Sickle Cell Disease," *Pain Management Nursing*, Vol. 17, No. 1, 2016, pp. 38–46.

Narayana, A., et al., “National Breakthrough Pain Study: Prevalence, Characteristics, and Associations with Health Outcomes,” *Pain*, Vol. 156, No. 2, 2015, pp. 252–259.

National Archives and Records Administration (NARA), “Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute-Care Hospitals and the Long-Term Care Hospital Prospective Payment System and Proposed Policy Changes and Fiscal Year 2018 Rates; Quality Reporting Requirements for Specific Providers; Medicare and Medicaid Electronic Health Record (EHR) Incentive Program Requirements for Eligible Hospitals, Critical Access Hospitals, and Eligible Professionals; Provider-Based Status of Indian Health Service and Tribal Facilities and Organizations; Costs Reporting and Provider Requirements; Agreement Termination Notices,” *Federal Register*, Vol. 82, No. 81, April 28, 2017, pp. 19796–20231. As of February 6, 2019:
<https://www.federalregister.gov/d/2017-07800>

———, “Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2018,” *Federal Register*, Vol. 82, No. 84, May 3, 2017, pp. 20690–20747. As of February 6, 2019:
<https://www.federalregister.gov/d/2017-08428>

———, “Prospective Payment System and Consolidated Billing for Skilled Nursing Facilities for FY 2018, SNF Value-Based Purchasing Program, SNF Quality Reporting Program, Survey Team Composition, and Proposal To Correct the Performance Period for the NHSN HCP Influenza Vaccination Immunization Reporting Measure in the ESRD QIP for PY 2020,” *Federal Register*, Vol. 82, No. 85, May 4, 2017, pp. 21014–21100. As of February 6, 2019:
<https://www.federalregister.gov/d/2017-08521>

———, “CY 2018 Home Health Prospective Payment System Rate Update and Proposed CY 2019 Case-Mix Adjustment Methodology Refinements; Home Health Value-Based Purchasing Model; and Home Health Quality Reporting Requirements,” *Federal Register*, Vol. 82, No. 144, July, 28, 2017, pp. 35270–35393. As of February 6, 2019:
<https://www.federalregister.gov/d/2017-15825>

Nygaard, I., et al., “Urinary Incontinence and Depression in Middle-Aged United States Women,” *Obstetrics and Gynecology*, Vol. 101, No. 1, 2003, pp. 149–156.

Pilkonis, P. A., et al., “Item Banks for Measuring Emotional Distress from the Patient-Reported Outcomes Measurement Information System (PROMIS(R)): Depression, Anxiety, and Anger,” *Assessment*, Vol. 18, No. 3, 2011, pp. 263–283.

- RAND Corporation, *Technical Expert Panel Summary/Expert Input Report: Development and Maintenance of Post-Acute Care Cross-Setting Standardized Patient Assessment Data*, Santa Monica, Calif.: RAND Corporation, WR-1187-CMMS, June 2016. As of February 6, 2019: https://www.rand.org/pubs/working_papers/WR1187.html
- Rapp, M. A., et al., “Cognitive Decline in Patients with Dementia as a Function of Depression,” *American Journal of Geriatric Psychiatry*, Vol. 19, No. 4, 2011, pp. 357–363.
- Rozich, J., and R. Roger, “Medication Safety: One Organization’s Approach to the Challenge,” *Journal of Clinical Outcomes Management*, Vol. 8, No. 10, 2001, pp. 27–34.
- Ruo, B., et al., “Depressive Symptoms and Health-Related Quality of Life: The Heart and Soul Study,” *Journal of the American Medical Association*, Vol. 290, No. 2, 2003, pp. 215–221.
- Saliba, D., and J. Buchanan, *Development and Validation of a Revised Nursing Home Assessment Tool: MDS 3.0*, Santa Monica, Calif.: RAND Corporation, Contract No. 500-00-0027/Task Order #2, April 2008. As of February 6, 2019: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/downloads/MDS30FinalReport.pdf>
- Schoevers, R., et al., “Association of Depression and Gender with Mortality in Old Age: Results from the Amsterdam Study of the Elderly (AMSTEL),” *British Journal of Psychiatry*, Vol. 177, No. 4, 2000, pp. 336–342.
- Seo, S., “A Review and Comparison of Methods for Detecting Outliers in Univariate Data Sets,” thesis, University of Pittsburgh, 2006.
- Shao, H., et al., “Diagnosed Depression Among Medicare Home Health Patients: National Prevalence Estimates and Key Characteristics,” *Psychiatric Services*, Vol. 62, No. 5, 2011, pp. 538–540.
- Smalbrugge, M., et al., “Pain Among Nursing Home Patients in the Netherlands: Prevalence, Course, Clinical Correlates, Recognition and Analgesic Treatment—An Observational Cohort Study,” *BMC Geriatrics*, Vol. 7, No. 3, 2007. As of February 6, 2019: <https://bmcgeriatr.biomedcentral.com/articles/10.1186/1471-2318-7-3>.
- Spitzer, R. L., et al., “Health-Related Quality of Life in Primary Care Patients with Mental Disorders: Results from the PRIME-MD 1000 Study,” *Journal of the American Medical Association*, Vol. 274, No. 19, 1995, pp. 1511–1517.
- Stein, C. R., R. B. Devore, and B. E. Wojcik, “Calculation of the Kappa Statistic for Inter-Rater Reliability: The Case Where Raters Can Select Multiple Responses from a Large Number of Categories,” *SUGI 30 Proceedings*, Philadelphia, Penn.: SAS Institute, Inc., April 2005. As of February 6, 2019: <https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/180-30.pdf>

- Teno, J. M., et al., "Daily Pain That Was Excruciating at Some Time in the Previous Week: Prevalence, Characteristics, and Outcomes in Nursing Home Residents," *Journal of the American Geriatrics Society*, Vol. 52, No. 5, 2004, pp. 762–767.
- Tukey, J. W., *Exploratory Data Analysis*, Reading, Mass.: Addison-Wesley Publishing Company, 1977.
- Uebersax, J., "Calculating Kappa with SAS," 2002. As of February 6, 2019:
<https://www.john-uebersax.com/stat/saskappa.htm>
- Unutzer, J., et al., "Long-Term Cost Effects of Collaborative Care for Late-Life Depression," *American Journal of Managed Care*, Vol. 14, No. 2, 2008, pp. 95–100.
- Van Almenkerk, S., et al., "Pain Among Institutionalized Stroke Patients and its Relation to Emotional Distress and Social Engagement," *International Journal of Geriatric Psychiatry*, Vol. 30, No. 10, 2014, pp. 1023–1031.
- Warshaw, E., et al., "Clinical and Cost Effectiveness of a Cleanser Protectant Lotion for Treatment of Perineal Skin Breakdown in Low-Risk Patients with Incontinence," *Ostomy/Wound Management*, Vol. 48, No. 6, 2002, pp. 44–51.
- Williams, J. W., et al., "Identifying Depression in Primary Care: A Literature Synthesis of Case-Finding Instruments," *General Hospital Psychiatry*, Vol. 24, No. 4, 2002, pp. 225–237.
- Xu, S., and M. F. Lorber, "Interrater Agreement Statistics with Skewed Data: Evaluation of Alternatives to Cohen's Kappa," *Journal of Consulting and Clinical Psychology*, Vol. 82, No. 6, 2014, pp. 1219–1227.
- Ziegelstein, R. C., "Depression in Patients Recovering from a Myocardial Infarction," *Journal of the American Medical Association*, Vol. 286, No. 13, 2001, pp. 1621–1627.