Development and Evaluation of Candidate Standardized Patient Assessment Data Elements: Findings from the National Beta Test (Volume 2: Background and Methods)

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

Prepared by
Maria Orlando E delen
Anthony Rodriguez
Sangeeta C. Ahluwalia
Emily K. Chen
Catherine C. Cohen
Sarah Dalton
Jason Michel Etchegaray
Wenjing Huang
Brenda Karkos
David J. Klein
Jaime Madrigano
Monique Martineau
Terry Moore
Patrick Orr
Jessica Phillips
Ben Schalet
Victoria Shier
Susan M. Paddock
Denis Agniel
Development and Evaluation of Candidate Standardized Patient Assessment Data Elements: Findings from the National Beta Test (Volume 2: Background and Methods)

by

Maria Orlando Edelen
Anthony Rodriguez
Sangeeta C. Ahluwalia
Emily K. Chen
Catherine C. Cohen
Sarah Dalton
Jason Michel Etchegaray
Wenjing Huang
Brenda Karkos
David J. Klein
Jaime Madrigano
Monique Martineau
Terry Moore
Patrick Orr
Jessica Phillips
Ben Schalet
Victoria Shier
Susan M. Paddock
Denis Agniel

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

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

This project was sponsored by the Centers for Medicare & Medicaid Services under contract no. HHSM-500-2013-13014I. The research was carried out within the Quality Measurement and Improvement Program in RAND Health Care, a division of the RAND Corporation. For more information, please see www.rand.org/health-care.

Published in 2019
Preface

The Centers for Medicare & Medicaid Services (CMS) contracted with the RAND Corporation to identify and develop standardized patient assessment data elements (SPADEs) for use in the following post-acute care (PAC) patient assessment instruments: the Outcome and Assessment Information Set, used in home health agencies; the Inpatient Rehabilitation Facility Patient Assessment Instrument, used in inpatient rehabilitation facilities; the Long-Term Care Hospital Continuity Assessment Record and Evaluation Data Set, used in long-term care hospitals; and the Minimum Data Set, used in nursing homes and skilled nursing facilities. RAND was tasked with developing and testing data elements within five areas of focus that fall under the clinical categories delineated in the Improving Medicare Post-Acute Care Transformation (IMPACT) 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 eight-volume report presents background information and results of the National Beta Test, which assessed a set of data elements within the five categories under the IMPACT Act. The National Beta Test was conducted between November 2017 and August 2018. Volume 1 is an executive summary of the material presented in the subsequent volumes. Volume 2 covers the data elements tested; the design; the sampling plan; information on training, recruitment, and retention; information on the data collection process; and the analytic plan. Volume 3 provides a sample description and reports analyses that evaluate the generalizability of results from the National Beta Test sample, both in terms of the representativeness of the facility/agency-level sample to the national population of PAC facilities/agencies, as well as the patients and residents who participated in the National Beta Test relative to the national population of patients and residents receiving PAC in the United States. Volumes 4–8 present the quantitative and qualitative data gathered during testing, as well as interpretations of the results for SPADEs in the following clinical categories: cognitive function (Volume 4), mental status and pain (Volume 5), impairments and special services, treatments, and interventions (Volume 6), and data elements that fall into other clinical categories (care preferences, medication reconciliation, and global health; Volume 7). Volume 8 describes the results and recommendations for SPADEs developed specifically for patients and residents who are unable to communicate (staff assessments of mental status, mood, and pain).
Contents

Preface ............................................................................................................................................ iii
Tables ............................................................................................................................................. vi
Abbreviations ................................................................................................................................ vii
1. Introduction ................................................................................................................................. 1
   Project Overview ........................................................................................................................... 1
   Previous Stakeholder Input Opportunities ..................................................................................... 2
2. Candidate SPADEs Included in National Beta Testing .............................................................. 5
   Selection and Development of Candidate SPADEs ........................................................................ 5
   Cognitive Function and Mental Status ............................................................................................ 5
   Pain .................................................................................................................................................. 9
   Impairments ................................................................................................................................ 10
   Special Services, Treatments, and Interventions .......................................................................... 11
   Other ............................................................................................................................................. 11
3. Design and Sampling ................................................................................................................ 13
   Design of Data Collection ............................................................................................................. 13
   Testing Targets ............................................................................................................................. 14
   Market and Facility Sampling ......................................................................................................... 16
4. Provider and Research Nurse Recruitment .............................................................................. 19
   Providers ..................................................................................................................................... 19
   Research Nurses ........................................................................................................................... 22
5. Training ................................................................................................................................... 24
   Research Nurses ............................................................................................................................ 24
   Facility/Agency Staff ...................................................................................................................... 25
   Refresher Trainings ......................................................................................................................... 27
   Replacement Training ..................................................................................................................... 27
6. Data Collection ......................................................................................................................... 29
   Field Management ......................................................................................................................... 29
   Data Management and Processing ................................................................................................ 31
   Data Reporting .............................................................................................................................. 31
   Data Delivery ................................................................................................................................. 31
   Additional Data Collection Activities ........................................................................................... 32
7. Analytic Plan ............................................................................................................................... 35
   Goals and Approach ...................................................................................................................... 35
Appendix A. Past Performance, Current Use, and Input Opportunities for Candidate
   SPADEs Tested in the National Beta Test ..................................................................................... 42
Appendix B. National Beta Test Research Nurse Training Agenda .............................................. 48
Appendix C. Example National Beta Test Field Training Agenda (Boston) .................................... 53
Tables

Table 3.1. National Beta Test Protocols .............................................................. 14
Table 3.2. Target Sample Sizes for National Beta Test Assessments for Communicative
   Admission, Communicative Discharge, and Non-Communicative, by PAC Setting .... 15
Table 4.1. Number of Facilities/Agencies Recruited and Subset That Participated in Training by
   Market ........................................................................................................... 22
Table 4.2. Number of Facilities/Agencies That Participated in Training by PAC Type ........ 22
Table 5.1. Facility/Agency Staff Training Schedule ............................................. 26
Table A.1. Past Performance, Current Use, and Input Opportunities for Candidate SPADEs
   Tested in the National Beta Test ................................................................. 42
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
DIF  differential item functioning
FY  fiscal year
HHA  home health agency
HRR  hospital referral region
IMPACT Act  Improving Medicare Post-Acute Care Transformation Act of 2014
IRF  inpatient rehabilitation facility
IRF-PAI  Inpatient Rehabilitation Facility Patient Assessment Instrument
IRT  item response theory
IV  intravenous
LCDS  Long-Term Care Hospital 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
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) (Pub. L. 113-185). 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 the Outcome and Assessment Information Set (OASIS) for HHAs, Inpatient Rehabilitation Facility Patient Assessment Instrument (PAI) for IRFs, Long-Term Care Hospital 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 distinct data elements 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 researchers 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
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 National 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 (FY) 2018/calendar year (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/pubs/working_papers/WR1187.html.1

---

1 RAND Corporation, 2017a.
• **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 PC 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/agency staff conducted 133 paired assessments so that results could be compared for both feasibility and interrater reliability. 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 PC 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 in the Federal Register at https://www.federalregister.gov/d/2017-07800 (for LTCHs); https://www.federalregister.gov/d/2017-08521 (for SNFs); https://www.federalregister.gov/d/2017-08428 (for IRFs); and https://www.federalregister.gov/d/2017-15825 (for HHAs).6

---

2 CMS, 2016.
3 Edelen et al., 2017.
4 RAND Corporation, 2017b.
5 CMS, 2018.
6 CMS, 2017a, 2017b, 2017c, 2017d.
• **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/agency staff conducted paired assessment (for communicative patients, 118 at admission and 42 at discharge; for non-communicative patients, 44 assessments) so that results could be compared for both feasibility and interrater reliability. The Alpha 2 feasibility test report is available at

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.

This report is the second volume in a series of eight reports that describe the methods and results from the National Beta Test. This second volume focuses on the design and methods for the National Beta Test.

---

7 Edelen et al., 2018.
2. Candidate SPADEs Included in National Beta Testing

The National Beta Test 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 National Beta Testing, and the activities through which quantitative and qualitative feedback were gathered on them.

Selection and Development of Candidate SPADEs

Candidate SPADEs in National 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 the National Beta Test can be found in Table A.1 in Appendix A. The table covers whether each group of data elements is currently in use and in which PAC assessments, what evidence exists for the feasibility and reliability or validity of the data elements (if any), and the input opportunities that have occurred as part of this effort leading up to the National 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 National Beta Test: the Brief Interview for Mental Status (BIMS), Confusion Assessment Method (CAM), and Staff Assessment of Mental Status.

Brief Interview for Mental Status

The 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 the 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 themselves 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 the individual’s ability to follow instructions, thereby posing a health and safety risk.

Two different versions of data elements assessing expression and understanding were included in National 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.

8 Adapted from Inouye et al., 1990. The Confusion Assessment Method (CAM) is the property of Hospital Elder Life Program. © 1988, 2003, Hospital Elder Life Program. All rights reserved.
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. 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. 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.

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. 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. It can also exacerbate other chronic medical conditions, compromise treatment participation and compliance, slow recovery from injuries and surgeries, and lead to rehospitalization. Although depression in the elderly often goes undiagnosed, prognosis is usually good when there is prompt recognition and treatment.

---

10 Dougherty et al., 1992.; Lequerica et al., 2007.
12 Banerjee and Macdonald, 1996; Bruce et al., 2002; Ell et al., 2006; Jones, Marcantonio, and Rabinowitz, 2003; Shao et al., 2011.
13 Lebowitz et al., 1997; Diefenbach, Tolin, and Gilliam, 2012; Kroenke et al., 2010; Spitzer et al., 1995; Harris and Cooper, 2006; Kane, Yochim, and Lichtenberg, 2010; Schoevers et al., 2000.
14 Ishii, Streim, and Saliba, 2010; Ishii, Streim, and Saliba, 2012; Rapp et al., 2011; Ziegelstein, 2001; Carson et al., 2003; DiMatteo, Lepper, and Croghan, 2000; Katon et al., 2004; Morley, 2007.
15 Alexopoulos et al., 2001; Areán and Cook, 2002.
Studies have also shown that treatment of depression in older adults can result in long-term cost savings.\textsuperscript{16} Four candidate SPADEs that assess mental status were included in National Beta Testing: Patient/Resident Mood Interview (Patient Health Questionnaire [PHQ]-2 to 9),\textsuperscript{17} 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).\textsuperscript{18}

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

The nine–data element 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-5).\textsuperscript{19} The PHQ-2 assesses the cardinal criteria for depression—depressed mood and anhedonia—using two data elements from the PHQ-9. The sensitivity of the PHQ-2, though acceptable, 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 data elements from the National Institutes of Health–supported PROMIS was identified for inclusion as possible SPADEs. The PROMIS data element bank for anxiety focuses on fear (e.g., fearfulness, feelings of panic), anxious misery (e.g., worry, dread), hyperarousal

\textsuperscript{16} Katon et al., 2008; Unützer et al., 2008.
\textsuperscript{17} Spitzer, Kroenke, and Williams, 1999. The Patient Health Questionnaire (PHQ) was developed by Pfizer Inc.© 1999 Pfizer Inc. All rights reserved.
\textsuperscript{18} Dewitt et al., 2018. PROMIS, *Patient-Reported Outcomes Measurement Information System*, and the PROMIS logo are marks owned by the U.S. Department of Health and Human Services.
\textsuperscript{19} American Psychiatric Association, 2013.
(e.g., tension, nervousness, restlessness), and somatic symptoms related to arousal (e.g., cardiovascular symptoms, dizziness). Because of the length of the complete data element bank and concerns that all data elements might not be relevant to the PAC population, we introduced a reduced list of data elements for consideration. PROMIS Anxiety was discussed in TEP 2, included in PC 2, and tested in Alpha 2.

PROMIS Depression
The PROMIS Depression data elements 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. Because of the length of the complete data element bank and concerns that all data elements might not be relevant to the PAC population, we introduced a reduced list of data elements for consideration. PROMIS Depression was discussed in a follow-up TEP webinar.

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. Pain is also associated with sleep disturbances, functional disability, and an increase in behavior problems, including agitation, irritability, and resistance to care, particularly for cognitively impaired patients. Even though pain is a common and recognizable human experience, it is often underrecognized, underdetected, and understudied among older adults. 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 National Beta Testing.

20 Pilkonis et al., 2011.
21 Amtmann et al., 2015; Wells, McCaffery, and Paseo, 2008; Husebo, Ballard, Fritze, et al., 2014; Iliffe et al., 2009.
22 McHugh, Casey, and Lawlor, 2011; Moscou-Jackson et al., 2016.
23 Lin et al., 2006; Narayana et al., 2015; Teno et al., 2004.
24 Husebo et al., 2014; Ishii et al., 2010; van Almenkerk et al., 2015.
25 Avenanti et al., 2005; Cadogan et al., 2005; Engle, Graney, and Chan, 2001; Ferrell, 2010; Jordan et al., 2011; Smalbrugge et al., 2007.
Pain Interview

Pain interview data elements included in National Beta Testing include assessment of presence of pain, frequency, severity, effect on sleep, interference with therapy- and nontherapy-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 National 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. 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 National 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. Incontinence is associated with a host of negative outcomes, including sleep difficulties, inactivity, social isolation, and depression. Changes in continence can signal 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. National Beta Testing included appliance use, frequency of incontinent events (chart review), and patient/resident perceived problem or burden with bladder and bowel incontinent events.

26 Ivers et al., 2000; Freeman et al., 2007.
27 Landefeld et al., 2008; Warshaw et al., 2002.
28 Nygaard et al., 2003; Landefeld et al., 2008.
(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 National Beta Testing. Nutritional approaches included parenteral/intravenous (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 National 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 occur during transition within, admission to, transfer to, or discharge from a hospital.29 At

---

29 Barnsteiner, 2005; Rozich and Resar, 2001; Gleason et al., 2004.
least 25 percent of all medication-related injuries are preventable.\textsuperscript{30} Medication reconciliation, 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 medication reconciliation 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.\textsuperscript{31} A standardized set of data elements that assess medication reconciliation 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 data elements included in National 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 the 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.

\textit{Global Health}

The goal of health-related quality of life 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 health-related quality of life through patient-reported outcomes 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. Patient-reported outcomes 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–supported PROMIS has developed data element banks for a large number of health-related quality of life domains using rigorous methodology.

The National Beta Test included PROMIS Global Health as a candidate SPADE. It consists of ten patient/resident interview data elements 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.

\textsuperscript{30} Institute of Medicine, 2007.
\textsuperscript{31} Joint Commission, 2015; Bao et al., 2012.
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 National 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 his or her market and to serve as a “gold standard” coder alongside a facility/agency staff person, allowing measurement of interrater reliability. This approach brought necessary robustness and consistency to both the testing and training processes. (For more information on training, see Chapter 5.)

To test performance of candidate SPADEs for different types of patients/residents and 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. **Non-communicative assessment**, to be administered at any point during a qualifying PAC stay among non-communicative patients/residents who were unable to make themselves understood in any fashion.
Table 3.1. National Beta Test Protocols

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Patient/Resident Eligibility</th>
<th>Assessment Windowa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicative admission</td>
<td>Admitted to PAC site under a new Medicare Part A or Part C covered stay; able to make themselves understood using any means</td>
<td>Day 1 (Admission Day) through Day 7</td>
</tr>
<tr>
<td>Communicative discharge</td>
<td>Any patient/resident for whom a National Beta Test admission assessment had been conducted</td>
<td>Day 1 (Discharge Day) through Day 3 (2 days prior to Discharge Day)</td>
</tr>
<tr>
<td>Non-communicative</td>
<td>Any patient/resident who is unable to make themselves understood using any means</td>
<td>48 hours, beginning on any day during the PAC stay the research nurse and field staff partner identify</td>
</tr>
</tbody>
</table>

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, PROMIS Anxiety, and PROMIS 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 other form. The 14 National Beta Test 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 National Beta Test are available online, 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.

---

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

<table>
<thead>
<tr>
<th>Assessment Type</th>
<th>LTCH (n = 28)</th>
<th>IRF (n = 28)</th>
<th>SNF (n = 84)</th>
<th>HHA (n = 70)</th>
<th>Total (n = 210)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicative admission</td>
<td>840</td>
<td>840</td>
<td>2,100</td>
<td>1,750</td>
<td>5,530</td>
</tr>
<tr>
<td>Communicative discharge</td>
<td>579</td>
<td>756</td>
<td>1,491</td>
<td>840</td>
<td>3,666</td>
</tr>
<tr>
<td>Non-communicative</td>
<td>280</td>
<td>280</td>
<td>420</td>
<td>350</td>
<td>1,330</td>
</tr>
</tbody>
</table>

The National 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 (targets of 840 in LTCHs, 840 in IRFs, 2,100 in SNFs, and 1,750 in HHAs). A subset of the admission assessments (targets 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 interrater reliability. Another subset of the admission assessments (targets 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 3, and then assessed again on a subset of the candidate SPADEs on admission Day 5 and admission Day 7. 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 National Beta Testing was estimated accordingly.

For non-communicative patients/residents, both facility/agency staff and research nurses administered the non-communicative 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 non-communicative 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: to detect 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 (Days 3, 5, and 7) 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, IRF-PAI, LCDS, and MDS 3.0) were also obtained from CMS for all patients/residents who participated in National Beta Testing. Data from these routine assessments were used to contribute to the evaluation of the validity of the candidate SPADEs. See Chapter 7 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 National 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 National Beta Test sample. The Denver, Colorado, HRR was eliminated from the National Beta Test 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. Because 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
- San Diego
- St. Louis.

**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 file (data version December 2016). The Provider of Services 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 CY 2014. These files included the number of Medicare discharges per year, which was important for identifying facilities of sufficiently large size to yield the minimum number of National Beta Test assessments. To be eligible, providers were required to have at least 72 Medicare discharges annually for LTCHs and IRFs, 60 stays annually per SNF, and 60 episodes for HHAs. 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 National 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. Because of 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
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.
4. Provider and Research Nurse Recruitment

With the plans for the testing design and sampling in place, the next step in setting up the National 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 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 and 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 National Beta Test goals and structure, data collection requirements, the vital role of participating providers, incentives for participation, how the data from the National 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 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 customer relations management 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 National 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)

A second phase of active recruitment was added to ensure a sufficient number of providers in the National Beta Test, 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 National 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 National Beta Test ultimately ended up withdrawing from the study. Generally, providers withdrew at this stage of the study because of 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 National 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

<table>
<thead>
<tr>
<th>Market</th>
<th>Recruited</th>
<th>Participated (% Recruited)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston</td>
<td>24</td>
<td>20 (83%)</td>
</tr>
<tr>
<td>Chicago</td>
<td>15</td>
<td>10 (67%)</td>
</tr>
<tr>
<td>Dallas</td>
<td>18</td>
<td>12 (67%)</td>
</tr>
<tr>
<td>Durham</td>
<td>16</td>
<td>13 (81%)</td>
</tr>
<tr>
<td>Fort Lauderdale</td>
<td>13</td>
<td>13 (100%)</td>
</tr>
<tr>
<td>Harrisburg</td>
<td>16</td>
<td>9 (56%)</td>
</tr>
<tr>
<td>Houston</td>
<td>17</td>
<td>16 (94%)</td>
</tr>
<tr>
<td>Kansas City</td>
<td>16</td>
<td>13 (81%)</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>22</td>
<td>17 (77%)</td>
</tr>
<tr>
<td>Nashville</td>
<td>18</td>
<td>14 (78%)</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>16</td>
<td>14 (88%)</td>
</tr>
<tr>
<td>Phoenix</td>
<td>14</td>
<td>11 (79%)</td>
</tr>
<tr>
<td>San Diego</td>
<td>19</td>
<td>10 (53%)</td>
</tr>
<tr>
<td>St. Louis</td>
<td>16</td>
<td>14 (88%)</td>
</tr>
<tr>
<td>Total</td>
<td>240</td>
<td>186 (78%)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>PAC Type</th>
<th>Participated</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHA</td>
<td>57</td>
</tr>
<tr>
<td>IRF</td>
<td>28</td>
</tr>
<tr>
<td>LTCH</td>
<td>28</td>
</tr>
<tr>
<td>SNF</td>
<td>73</td>
</tr>
</tbody>
</table>

Research Nurses

The National 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 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 National 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.
5. Training

Data collection during the National Beta Test 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 National 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 National Beta Test 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 National Beta Test 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 National Beta Test 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 National Beta Test assessments.

Pretraining Webinar

This two-hour webinar, held the week before the in-person training, provided essential context for the study, for the National Beta Test, and for the research nurses’ role as trainers and
“gold standard” data collectors. Specific topics included an overview of the National 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 National Beta Test assessment form and accompanying user manual, for offline study prior to the in-person training; and a moderated question-and-answer session and review of training logistics.

**In-Person Research Nurse Training**

The in-person National Beta Test 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 National Beta Test 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 National Beta Test 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 National Beta Test 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 National Beta Test assessment with feedback from trainers, and the training team to troubleshoot any issues associated with data collection tablets.

Although 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 might have replaced staff trained during Phase 1. The first phase 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**

<table>
<thead>
<tr>
<th>Phase</th>
<th>Training Dates</th>
<th>West Region</th>
<th>Central Region</th>
<th>East Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1</td>
<td>October 9 and 10</td>
<td>Los Angeles</td>
<td></td>
<td>Boston</td>
</tr>
<tr>
<td></td>
<td>October 18 and 19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>October 23 and 24</td>
<td>Los Angeles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>September 25 and 26</td>
<td>San Diego</td>
<td>Kansas City</td>
<td>Philadelphia</td>
<td></td>
</tr>
<tr>
<td>November 8 and 9</td>
<td>Phoenix</td>
<td>St. Louis</td>
<td>Harrisburg</td>
<td></td>
</tr>
<tr>
<td>November 15 and 16</td>
<td>Dallas</td>
<td>Nashville</td>
<td>Durham</td>
<td></td>
</tr>
<tr>
<td>November 29 and 30</td>
<td>Houston</td>
<td>Chicago</td>
<td>Fort Lauderdale</td>
<td></td>
</tr>
<tr>
<td>Phase 2</td>
<td>January 30 and 31</td>
<td>San Diego</td>
<td>Chicago</td>
<td>Fort Lauderdale</td>
</tr>
</tbody>
</table>

26
Refresher Trainings

Because of the lengthy data collection field period (i.e., approximately six months with an extended period from 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 the following refresher training to both research nurses and facility/agency staff:

- **Refresher Training Webinars (January 11, 16, and 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 National 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 questions 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 3, Day 5, and Day 7 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 because of illness, two were dismissed for performance issues, and one after misunderstanding job expectations. Qualidigm re-recruited and replaced research nurses lost because of 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.
6. Data Collection

National Beta Test data collection began immediately after training and practice assessments were complete. Because facility/agency 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 National Beta Test 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

Because of the complex data collection design and procedures and 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 National Beta Test 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 National 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 National 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 and 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 National 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 National Beta Test providers.
- **Provider webinars**: We held a series of webinars during the data collection phase to emphasize the role and importance of National 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 National 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 of all downloaded assessments. This multistep quality assurance process first entailed verifying that all key variables that tracked site, survey type (research nurse or facility/agency staff assessment), assessment type (admission communicative, non-communicative 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 or 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/agency 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 quality assurance 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 National 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 Volumes 4–8.

**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 HHA 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 Quality Indicator inbox, during Special Open Door Forums). Specifically, the survey was designed to collect information on perspectives of the clinical utility of National Beta Test 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 National 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 data elements, each data element was ranked relative to other items within the SPADE. Ranked data elements were displayed with stacked bar charts.
Facility/Agency Staff Focus Groups

Facility/agency staff who participated in the National Beta Test hold valuable 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 National 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 National Beta Test data elements were particularly easy or difficult to complete, which data elements were most clinically relevant and useful across PAC settings, and how the facility/agency staff would like to see these data elements 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 who 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 National Beta Test, such as time to complete particular candidate SPADEs, warranted follow-up. Similarly, probes were generated to get additional information about select results of the National Beta Test 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 valuable perspectives on data collection during the National 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 and to potentially inform future clinical testing by CMS. Specifically, RAND sought to gather research nurses’ understanding of whether the concepts measured by National Beta Test data elements 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 who 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.
7. Analytic Plan

Goals and Approach

For all communicative and non-communicative admission assessments, there were four primary goals for the data analysis: (1) to compute descriptive statistics for candidate SPADEs, (2) to determine the feasibility of administration, (3) to evaluate interrater reliability between pairs of research nurses and facility/agency staff assessors, and (4) to determine the content validity of candidate SPADEs and distribution of scores. Goals 1 and 2 applied also to communicative discharge assessments.

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

Many of the analyses conducted during National 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 researchers 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), length of stay, clinical characteristics (e.g., sepsis, heart failure, stroke), discharge disposition, and ability to perform two activities of daily living (i.e., toileting [hygiene] and ability to transfer from lying to sitting [mobility]). For provider characteristics, we considered, when appropriate, differences by whether a facility/agency is located in an urban or nonurban area, region of the country (Northeast, South, Midwest, West), facility ownership (for-profit versus nonprofit) and facility size (above or below setting-type median).

Descriptive Statistics

Response frequencies were computed for each candidate SPADE using data collected from facility/agency staff and unpaired research nurse admission and discharge assessments to provide descriptive information pertaining to category response distributions overall and by PAC setting. Additionally, for select candidate SPADEs (e.g., evidence of change in mental status), basic
descriptive statistics were computed for patient/resident characteristics (e.g., demographics, clinical subgroups) overall and by PAC setting.

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 and 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 data element 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 data element had missing data or a category response was endorsed, indicating an inability to assess or no answer (e.g., “Unable to assess/no response”).

Completion Time

Completion times were estimated based on the subset of assessments completed by facility/agency staff, because they are the primary group of interest for evaluating the ease of administration. Time spent to complete the data elements was self-reported by facility/agency staff and collected via time stamps recorded on tablets during the assessment administration. These time sources were both considered in order to maximize available data. 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 times were available, the self-report estimate was preferred unless it was clearly out of range (e.g., 140 minutes), in which case the tablet time was used. The final data set was inspected for outliers to be excluded. Values that fell more than 1.5 times the interquartile range above the third quartile were excluded from time estimate calculations, using an established method. Time to complete each candidate SPADE was evaluated overall and by PAC setting as well as by provider characteristics (e.g., urban versus nonurban). Completion times were compared across these groupings and statistically evaluated using analysis of variance methods.

---

31 Seo, 2006; Tukey, 1977.
**Interrater Reliability**

To determine whether data elements could be completed with acceptable interrater reliability, we calculated the level of agreement between pairs of research nurses and facility/agency staff assessors. For dichotomous data, we computed kappa and overall agreement percentage. For ordinal data element data, we computed weighted kappa, with Fleiss-Cohen (quadratic) weights. For continuous/approximately continuous data (more than five ordered categories), we computed Pearson’s correlations. 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 pseudo-cases for the missing response categories are created to produce a square table that allows for kappa to be computed. These pseudo-cases are given infinitesimally small weights so as to have no effect on the kappa coefficient.\(^{34}\)

For all data elements, we report raw percent agreement as a supplemental or alternative index of interrater reliability. 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.\(^{35}\) As such, the analytic team compared observed prevalence rates for all data elements with 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, as determined by study power calculations, we emphasized raw percent agreement as a measure of interrater reliability. Evaluative labels for different interrater reliability ranges are conventionally used to interpret interrater reliability results (slight/poor agreement = 0 to 0.20, 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).

Interrater reliability coefficients were computed overall, for each PAC setting type, and by provider characteristics. Because interrater reliability results are highly dependent on the prevalence of a given symptom/item response in the population being assessed (a.k.a. between-subjects variance), interrater reliability coefficients were calculated for subpopulations but not directly compared across those subpopulations (e.g., PAC settings, urban versus nonurban), where this prevalence is likely to differ. Thus, interpretation of subpopulation interrater reliability values were made for each subpopulation and in the context of examining and reporting of prevalence and raw rater agreement.

**Validity**

The focus of goal 4 was on the convergent validity of candidate SPADEs. That is, we focused on examining associations between National Beta Test data elements that should be

---

\(^{34}\) Stein, Devore, and Wojcik, 2005; Uebersax, 2002.

\(^{35}\) Cicchetti and Feinstein, 1990; Xu and Lorber, 2014; Byrt, Bishop, and Carlin, 1993; McHugh, 2002.
theoretically related based on literature and clinical practice. These associations were examined for each candidate SPADE, but ultimately analyses were prioritized for SPADEs 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 the presence or absence of select clinical diagnoses. Goal 4 focused on evaluating and comparing frequency and score distributions by these identified groups. Merging National Beta Test and legacy assessment data allowed us to identify conditions (e.g., sepsis, heart failure, 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 National Beta Test and legacy instrument data element. For ordinal data elements, 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) data elements were conducted using a standard chi-squared test.

Chart Review Look-Back Evaluation

For chart review look-back evaluation (admission and discharge) candidate SPADEs, the goal was to understand which look-back time frame (i.e., on admission [Day 1], and on Days 3, 5, and 7; Discharge Day and Discharge Day minus 2) would be most appropriate for these data elements. 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 purposes. Frequencies across Days (1, 3, 5, 7) were converted to cumulative percentages to observe rates of increase. Like admission data, discharge data frequency distributions were converted to cumulative percentages to reflect percent occurrence on Discharge Day minus 2 and on Discharge Day. 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. If rate of occurrence did not depend on look-back period (e.g., the vast majority of occurrences are noted on admission [Day 1]), then data across look-back days were
collapsed to reflect “noted” versus “not noted.” Analyses were done overall and by PAC setting type.

Although it is possible that each candidate SPADE offers an ideal look-back option, in which the greatest information is captured within a shortest look-back period, we were primarily interested in an overall best look-back option for all candidate SPADEs. To that end, if there was variability in ideal look-backs provided across candidate SPADE, then recommendations were made using the most frequent look-back supported by most data elements. However, if occurrence did not depend on look-back (that is, nothing or very little was gained by a longer look-back), then no recommendation was made as the data remain relatively unaffected by different look-back periods (e.g., same information on Day 1 as Day 5).

Repeat Assessment Evaluation

For repeat assessments, the goal was to identify which assessment days (i.e., Day 3, 5, 7) are most appropriate for these data elements. 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 frequencies for type of change (e.g., no change, improve) across assessment days.

To statistically compare observed differences on continuous data elements (or ordinal data elements with five or more ordered categories) over repeat assessments, within-subjects analysis of variance followed by pairwise comparisons were tested. 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 data elements 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 (e.g., improve, decline). These results are discussed to inform 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 whether there were significant statistical differences and offer possible explanations for such differences. Moreover, when appropriate, the implications of administering candidate SPADEs on different assessment days are discussed.

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 and by PAC setting to determine whether observed overall differences varied by setting.

Alternate Forms

For alternate forms, goal 6 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 data elements, 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 the National Beta Test (two–data element LCDS/IRF-PAI version and the three–data element MDS version) assess patients’/residents’ ability to express themselves and understand others. However, a key distinction is that the two–data element version conflates expression with speech clarity, but the three–data element version separates these. Therefore, the goal was to evaluate and compare data collected on similar data elements from each version and to assess the added utility of the speech clarity data element.

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

Psychometric Properties

Goal 7 was to understand the psychometric properties of scalable data elements (e.g., PROMIS Anxiety). As an initial step, we examined basic data element-level and scale reliability statistics. This included the percentage of responses at the floor and ceiling, skew, kurtosis, average inter–data element correlations, data element–total correlations, and Cronbach’s alpha. These statistics are used to identify poorly functioning data elements 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 data element–total correlations and explained common variance. Low values of data element–total correlations (less than 0.40) may indicate problematic data elements. For explained common variance, values greater than or equal to 0.60 may indicate that a data element set is sufficiently unidimensional. Data elements that form a sufficiently unidimensional scale were fit to the graded response model. The graded response model can be used to estimate the discrimination and location parameter(s) of the data elements that are discussed.

Differential item functioning (DIF) was also examined. DIF occurs when data elements 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 (≤ 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-R² measure of magnitude (≤ 0.02) in model comparisons. This
0.02 value for R² magnitude is conventionally used to identify nontrivial DIF in the development
of PROMIS instruments.
Appendix A. Past Performance, Current Use, and Input Opportunities for Candidate SPADEs Tested in the National Beta Test

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

<table>
<thead>
<tr>
<th>Data Element (Information Source)</th>
<th>Past Performance</th>
<th>Current Use in PAC Settings</th>
<th>Input Opportunities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive Function</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIMS (patient/resident interview)</td>
<td>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. 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.</td>
<td>The BIMS is currently collected in the MDS 3.0 and the IRF-PAI.</td>
<td>TEP 1, PC 1, proposed rulemaking</td>
</tr>
<tr>
<td>CAM (multiple information sources)</td>
<td>The Short CAM has been shown to be effective in identifying delirium in validated research studies. 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). 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).</td>
<td>MDS 3.0 and LCDS currently use versions of the four-item CAM (Short CAM), but response options and wording differ slightly.</td>
<td>TEP 1, PC 1, proposed rulemaking</td>
</tr>
<tr>
<td>Staff Assessment of Mental Status (observation)</td>
<td>Studies testing in nursing home patients have shown it to have moderate interrater reliability ($r = 0.80$) and good validity based on its correlation with other assessments such as the Blessed Test ($r = 0.66, p &lt; 0.05$) and the Reisberg Global Deterioration Scale ($r = 0.59, p &lt; 0.05$).</td>
<td>Included in the MDS 3.0.</td>
<td>TEP 2, PC 2, Alpha 2</td>
</tr>
<tr>
<td>Data Element (Information Source)</td>
<td>Past Performance</td>
<td>Current Use in PAC Settings</td>
<td>Input Opportunities</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------------</td>
<td>-----------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td><strong>Cognitive Function</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expression and Understanding (multiple information sources)</td>
<td>Similar data elements were tested in the PAC-PRD and formed a composite <em>Communication</em> variable, which had substantial agreement for interrater reliability (kappa range of 0.74 to 0.80).&lt;sup&gt;g&lt;/sup&gt;</td>
<td>Speech Clarity, Makes Self Understood, and Ability to Understand Others are currently included in the MDS 3.0.</td>
<td>PC 1</td>
</tr>
<tr>
<td></td>
<td>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).&lt;sup&gt;h&lt;/sup&gt;</td>
<td>Expression of Ideas and Wants and Understanding Verbal Content are currently used in IRF-PAI and LCDS. OASIS uses a similar item, Speech and Oral (Verbal) Expression of Language, which differs in response option phrasing but uses similar tiers of difficulty.</td>
<td></td>
</tr>
<tr>
<td>Behavioral Signs and Symptoms: Presence and Frequency; Impact on Patient/Resident; Impact on Others; Rejection of Care (multiple information sources)</td>
<td>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.</td>
<td>The Behavioral Signs and Symptoms data elements are currently in the MDS 3.0. OASIS includes a similar data element, which records the frequency of disruptive behaviors.</td>
<td>TEP 2, PC 2, Alpha 2, proposed rulemaking</td>
</tr>
<tr>
<td><strong>Mental Status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient/Resident Mood Interview (PHQ-2 to 9; patient/resident interview)</td>
<td>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 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 National Beta Testing across the four PAC settings (kappas ranged from 0.74 to 0.91).&lt;sup&gt;i&lt;/sup&gt;</td>
<td>The PHQ-9 and PHQ-2 are currently implemented in the MDS 3.0 and OASIS, respectively.</td>
<td>TEP 1, PC 1, Alpha 1, proposed rulemaking (PHQ-2)</td>
</tr>
<tr>
<td>Staff Assessment of Patient/Resident Mood (PHQ-9-OV; observation)</td>
<td>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.&lt;sup&gt;j&lt;/sup&gt;</td>
<td>Currently included in the MDS 3.0.</td>
<td>TEP 2, PC 2, Alpha 2</td>
</tr>
<tr>
<td>Data Element (Information Source)</td>
<td>Past Performance</td>
<td>Current Use in PAC Settings</td>
<td>Input Opportunities</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>PROMIS Anxiety (patient/resident interview)</td>
<td>PROMIS item banks have been developed for a large number of topics using rigorous methodology. Eleven anxiety items from the PROMIS item bank were tested in Alpha 2 feasibility testing. Interrater reliability ranged from 0.80 to 1.00, with most ratings falling between 0.95 to 1.00, indicating almost perfect agreement between assessors. 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 seven-day look-back period.</td>
<td>PROMIS Anxiety items are not included in the current PAC assessments.</td>
<td>TEP 2, PC 2, Alpha 2</td>
</tr>
<tr>
<td>PROMIS Depression (patient/resident interview)</td>
<td>Because it was developed relatively recently, PROMIS Depression has not been as extensively validated as the other data elements included in National Beta Testing; it has, however, been shown to be sensitive to change in diverse clinical samples.</td>
<td>PROMIS Depression items are not included in the current PAC assessments.</td>
<td>Follow-up TEP webinar</td>
</tr>
<tr>
<td>Medical Conditions: Pain</td>
<td>The data elements related to effect of pain on sleep and daily activities have been shown to enhance the evaluation of pain. The data element cluster in the MDS has demonstrated excellent reliability, with kappas of 0.96 and 0.97. 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. 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.</td>
<td>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. OASIS asks about standardized pain assessment being conducted, and indication of severe pain. The IRF-PAI and LCDS have assessments of pain severity.</td>
<td>TEP 1, PC 1, Alpha 1, TEP 2, PC 2</td>
</tr>
<tr>
<td>Data Element (Information Source)</td>
<td>Past Performance</td>
<td>Current Use in PAC Settings</td>
<td>Input Opportunities</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Staff assessment of pain or distress (observation)</td>
<td>The MDS data element cluster has demonstrated excellent reliability, with kappas of 0.94 and 0.96. Interrater reliability in Alpha 2 feasibility testing ranged from 0.69 to 1.00, indicating substantial to almost perfect agreement for the items. Assessors’ feedback reflected that the pain items were straightforward but somewhat challenging to administer because of the time required for observation and the need to consult multiple data sources.</td>
<td>Observational data elements to assess pain are included in the OASIS and the MDS 3.0. The OASIS item queries the frequency of pain interfering with a patient’s/resident’s activity or movement.</td>
<td>TEP 1, TEP 2, PC 2, Alpha 2</td>
</tr>
</tbody>
</table>

**Impairments**

<table>
<thead>
<tr>
<th>Impairments</th>
<th>Past Performance</th>
<th>Current Use in PAC Settings</th>
<th>Input Opportunities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hearing, Vision (multiple information sources)</td>
<td>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). In national MDS 3.0 testing, these data elements showed almost perfect interrater reliability (kappa range of 0.86–0.94).</td>
<td>MDS 3.0 currently assesses the hearing and vision data elements that are included in National Beta Testing. OASIS also assesses hearing and vision abilities.</td>
<td>TEP 1, PC 1, proposed rulemaking</td>
</tr>
<tr>
<td>Bladder and Bowel Continence patient/resident perceived problem (patient/resident interview)</td>
<td>In Alpha 1 testing, the items had almost perfect interrater reliability with kappas of 0.90 or higher.</td>
<td>Data elements to assess patient/resident perceived problem or burden with bladder and bowel incontinent events are not currently collected in PAC assessments.</td>
<td>Alpha 1, PC 2</td>
</tr>
<tr>
<td>Data Element (Information Source)</td>
<td>Past Performance</td>
<td>Current Use in PAC Settings</td>
<td>Input Opportunities</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------------</td>
<td>-----------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Bladder and Bowel Continence Appliance use, frequency of events, need for assistance (chart review)</td>
<td>In the national MDS 3.0 testing, the bladder and bowel data element cluster had almost perfect interrater reliability (kappa of 0.95). Similar items were also tested in PAC-PRD with kappas ranging from 0.60 to 0.90. Alpha 1 results suggest that items performed sufficiently well across settings. Interrater reliability on bladder and bowel items was substantial to almost perfect for items that could be assessed (kappas ranged from 0.61 to 0.90). Some items did not have sufficiently complete information, and kappas could not be calculated.</td>
<td>Similar items to those tested in National 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).</td>
<td>TEP 1, Alpha 1, PC 2</td>
</tr>
</tbody>
</table>

**Special Services, Treatments, and Interventions (SSTIs)**

| Nutritional Approaches: Parental/IV Feeding, Feeding Tube, Mechanically Altered Diet, Therapeutic Diet (chart review) | In the national MDS 3.0 testing, the nutritional approaches items exhibited almost perfect interrater reliability (kappa range of 0.89 to 0.96). | The nutritional approaches items are currently collected in MDS 3.0. | 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) | A subset of the SSTIs (and without the subitems) was tested in PAC-PRD. Although interrater reliability was not calculated, the use supports the feasibility of assessing SSTIs across PAC settings. In national MDS 3.0 testing, a slightly different suite of SSTIs showed almost perfect interrater reliability (kappa of 0.84 and 0.90). | The SSTIs (except for IV access and the subitems) are currently included in MDS 3.0. | PC 1, TEP 2, proposed rulemaking |

**Other**

<p>| Care Preferences: Decisionmaking preferences, designated health care agent (multiple information sources) | In the national MDS 3.0 testing, this item had almost perfect agreement (kappa of 0.99). 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. The Advance Directive data element was tested in the PAC-PRD and found to be feasible to implement with a kappa of 0.72. | The MDS assesses the patient’s/resident’s preferences for family/significant other involvement in care discussions. | TEP 1, Alpha 1, TEP 2, PC 2 |</p>
<table>
<thead>
<tr>
<th>Data Element (Information Source)</th>
<th>Past Performance</th>
<th>Current Use in PAC Settings</th>
<th>Input Opportunities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication Reconciliation (multiple information sources)</td>
<td>Medication reconciliation items tested in Alpha 1 had the lowest interrater reliability of any data elements tested (kappa coefficients ranged from 0.06 to 0.83), and assessors experienced difficulties completing the assessment. Therefore, medication reconciliation showed considerable difficulty with reliability and feasibility. 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. Based on Alpha 2 feedback, the burden of the medication reconciliation items was reduced by eliminating the seven-day look-back period for the items on which medications were being taken, reducing the number of drug classes to six, and revising wording.</td>
<td>A similar item noting the classes of medications a patient/resident is taking is included in the MDS 3.0. However, none of the other medication reconciliation items tested in National Beta Testing are in the current PAC assessments.</td>
<td>TEP 1, Alpha 1, TEP 2, PC 2, Alpha 2</td>
</tr>
<tr>
<td>PROMIS Global Health (patient/resident interview)</td>
<td>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.</td>
<td>PROMIS Global Health is not included in any current PAC assessments but is used on several large national and state surveys.</td>
<td>PC 2</td>
</tr>
</tbody>
</table>

**SOURCES:**

a Gage, Smith, et al., 2012.
b Inouye et al., 2014.
c Saliba, and Buchanan, 2008.
d 2012.
e Casten et al., 1998.
f Lawton et al., 1998.
g Gage, Ingber, et al., 2012.
h Saliba and Buchanan, 2008.
i Klapow et al., 2002; Kroenke, Spitzer, and Williams, 2001; Löwe et al., 2004; Ruo et al., 2003; Williams et al., 2002.
j Saliba and Buchanan, 2008.
Appendix B. National Beta Test 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  
Sangeeta Ahluwalia, PhD
9:15 am  Field Staff Training Plan: Schedule and Roles  
Jason Etchegaray, PhD
9:30 am  National Beta Test Data Collection Procedures & Materials  
Sangeeta Ahluwalia, PhD
10:30 am  Break
10:45 am  Tablet Orientation and Walk-Through  
Jessica Phillips, MS, and Programming Team

Communicative Assessment: Patient/Resident Interview Data Elements

11:45 am  Data Element Training: Hearing, Vision, Expression, and Understanding  
Lisa Newton, MSN, RN-BC, NEA-BC
12:15 pm  Meet and Greet Luncheon
1:45 pm  Data Element Training: Cognition – Brief Interview for Mental Status (BIMS)  
Cindy Lippman, RN
2:15 pm  Data Element Training: PROMIS Global Health  
Maria Edelen, PhD
3:00 pm  Break
3:15 pm  Data Element Training: Pain Interview  
Lisa Newton, MSN, RN-BC, NEA-BC
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  
Robin Richardson, RN

9:30 am  Data Element Training: Mood – PROMIS Depression; Anxiety  
Maria Edelen, PhD

10:00 am  Data Element Training: Care Preferences Interview  
Julia Portale, MBA, MPH, LMSW

10:15 am  Data Element Training: Continence Interview  
Vicki Shier, PhD

10:45 am  Break

Communicative Assessment: Multiple Information Sources Data Elements

11:00 am  Data Element Training: Cognition – Confusion Assessment Method (CAM)  
Ann Spenard, DNP, RN-BC

11:30 am  Data Element Training: Behavioral Signs and Symptoms  
Ann Spenard, DNP, RN-BC

12:15 pm  Lunch

1:15 pm  Data Element Training: Medication Reconciliation  
Robin Richardson, RN

2:00 pm  Practice Exercise: Medication Reconciliation

2:45 pm  Data Element Training: Care Preferences  
Julia Portale, MBA, MPH, LMSW

3:00 pm  Break

3:15 pm  Data Element Training: Continence  
Vicki Shier, PhD

4:15 pm  Data Element Training: Special Services, Treatments, & Interventions (SSTI)  
Ann Spenard, DNP, RN-BC

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  
Facilitators: Alpha 1 & Alpha 2 Research Nurses

9:30 am  
Role Play Introduction & Schedule Review  
Sarah Dalton, MA

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)  
Brenda Karkos, MSN/MBA, RN, CHPN

9:15 am  Data Element Training: Pain  
Brenda Karkos, MSN/MBA, RN, CHPN

9:45 am  Data Element Training: Mood – PHQ-OV  
Robin Richardson, RN

10:15 am  Break

10:30 am  Non-communicative assessment modeling: Mood  
Robin Richardson, RN, and Cindy Lippman, RN

11:00 am  Review and Reinforcement: Core assessment processes  
Sangeeta Ahluwalia, PhD, and Jessica Phillips, MS

11:30 am  Field Data Collection “Lessons Learned” Panel  
Facilitators: Alpha 1 & Alpha 2 Research Nurses

12:15 pm  Lunch

1:15 pm  Field Staff Training Plan: Schedule, Assignments, & Logistics  
Jason Etchegaray, PhD

1:45 pm  Training and Presenter Tips  
Julia Rollison, PhD, MPH, and Lisa Newton, MSN, RN-BC, NEA-BC

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  Principal Trainer Meeting
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 (Continence 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 (Continence; 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 National Beta Test Field Training Agenda (Boston)

**Wednesday, October 18**

8:00 am  Sign-in, pick-up materials including tablet assignment

8:30 am  Welcome and Introductions
         Review of Training Agenda and Activities
         Emily Butcher, Regional Training Lead, RAND Corporation

         Data Collection Overview
         Patrick Orr, Regional Training Lead, RAND Corporation

9:30 am  Tablet Orientation & Modeling
         Emily Butcher, Regional Training Lead, RAND Corporation

10:30 am Break

**Communicative Assessment: Patient/Resident Interview Data Elements**

10:40 am  Hearing, Vision, Expression and Understanding
          Lynda Sheehan, RN

11:10 am  Cognition
          Brief Interview for Mental Status (BIMS)
          Ann Spenard, DNP, RN-BC

11:35 am  PROMIS
          Global Health
          Lynn Averill, RN

12:00 pm  Pain
          Brenda Karkos, MSN/MBA, RN, CHPN

12:30 pm  Lunch

1:15 pm  Mood
         PHQ2 to 9
         PROMIS Depression
         PROMIS Anxiety
         Ann Spenard, DNP, RN-BC
2:15 pm  Care Preferences
Brenda Karkos, MSN/MBA, RN, CHPN

2:30 pm  Continence
Lynda Sheehan, RN

2:45 pm  Break

2:55 pm  Tablet Reinforcement
Emily Butcher, Regional Training Lead, RAND Corporation

3:25 pm  Role play (in pairs) #1
Patrick Orr, Regional Training Lead, RAND Corporation

4:15 pm  Debrief Role play/Q&A
Patrick Orr, Regional Training Lead, RAND Corporation

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  
Brenda Karkos, MSN/MBA, RN, CHPN

2:00 pm  Pain  
Brenda Karkos, MSN/MBA, RN, CHPN

2:30 pm  Mood  
Staff Assessment of Patient/Resident Mood (PHQ-9-OV)  
Ann Spenard, DNP, RN-BC

3:00 pm  Break

3:10 pm  Role play and debrief (in pairs) # 2  
Patrick Orr, Regional Training Lead, RAND Corporation

3:50 pm  Role play and debrief (in pairs) # 3  
Patrick Orr, Regional Training Lead, RAND Corporation

4:15 pm  Q&A  
Next Steps  
Emily Butcher, Regional Training Lead, RAND Corporation
Appendix D. National Beta Test 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?
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.

<table>
<thead>
<tr>
<th>Rank (1-6)</th>
<th>a) Perceived Problem or Burden with Bladder Incontinent Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>b) Bladder Appliance Use</td>
<td></td>
</tr>
<tr>
<td>c) Bladder Appliance Use – Current Setting</td>
<td></td>
</tr>
<tr>
<td>d) Bladder Appliance Use – Primary Reason for Catheter Placement</td>
<td></td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Rank (1-5)</th>
<th>e) Bladder Appliance Use – Need for Assistance in Appliance Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>f) Bladder – Frequency of Incontinent Events</td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Rank (1-7)</th>
<th>a) Perceived Problem or Burden with Bowel Incontinent Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>b) Bowel Appliance Use</td>
<td></td>
</tr>
<tr>
<td>c) Bowel Appliance Use – Current Setting</td>
<td></td>
</tr>
<tr>
<td>d) Bowel Appliance Use – Need for Assistance in Appliance Management</td>
<td></td>
</tr>
<tr>
<td>e) Bowel – Frequency of Incontinent Events</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Not at all useful</th>
<th>Slightly useful</th>
<th>Somewhat useful</th>
<th>Moderately useful</th>
<th>Extremely useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Chemotherapy</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>b) Radiation</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>c) Oxygen therapy</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>d) Suctioning</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>e) Tracheostomy care</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>f) Invasive mechanical ventilator</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>g) Non-invasive mechanical ventilator</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
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?

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

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

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?

<table>
<thead>
<tr>
<th>Data Element</th>
<th>Not at all burdensome</th>
<th>Slightly burdensome</th>
<th>Somewhat burdensome</th>
<th>Moderately burdensome</th>
<th>Extremely burdensome</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) BIMS</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>b) Pain Interview</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>c) PHQ-2 to 9 interview</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>d) PROMIS Depression</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>e) Anxiety</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>f) Care Preferences</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>g) Continence interview</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>h) PROMIS Global Health</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Item</th>
<th>Rank (1-6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Perceived Problem or Burden with Bladder Incontinent Events</td>
<td></td>
</tr>
<tr>
<td>b) Bladder Appliance Use</td>
<td></td>
</tr>
<tr>
<td>c) Bladder Appliance Use – Current Setting</td>
<td></td>
</tr>
<tr>
<td>d) Bladder Appliance Use – Primary Reason for Catheter Placement</td>
<td></td>
</tr>
<tr>
<td>e) Bladder Appliance Use – Need for Assistance in Appliance Management</td>
<td></td>
</tr>
<tr>
<td>f) Bladder – Frequency of Incontinent Events</td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Item</th>
<th>Rank (1-5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Perceived Problem or Burden with Bowel Incontinent Events</td>
<td></td>
</tr>
<tr>
<td>b) Bowel Appliance Use</td>
<td></td>
</tr>
<tr>
<td>c) Bowel Appliance Use – Current Setting</td>
<td></td>
</tr>
<tr>
<td>d) Bowel Appliance Use – Need for Assistance in Appliance Management</td>
<td></td>
</tr>
<tr>
<td>e) Bowel – Frequency of Incontinent Events</td>
<td></td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Rank (1-7)</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>a)</td>
<td>Pain Presence</td>
</tr>
<tr>
<td>b)</td>
<td>Pain Frequency</td>
</tr>
<tr>
<td>c)</td>
<td>Pain Effect on Sleep</td>
</tr>
<tr>
<td>d)</td>
<td>Pain Interference, Therapy Activities</td>
</tr>
<tr>
<td>e)</td>
<td>Pain Interference, Other Activities</td>
</tr>
<tr>
<td>f)</td>
<td>Pain Severity – Numeric Pain Scale</td>
</tr>
<tr>
<td>g)</td>
<td>Pain Relief</td>
</tr>
</tbody>
</table>

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?

<table>
<thead>
<tr>
<th>Item</th>
<th>Not at all burdensome</th>
<th>Slightly burdensome</th>
<th>Somewhat burdensome</th>
<th>Moderately burdensome</th>
<th>Extremely burdensome</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Chemotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) Radiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) Oxygen therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) Suctioning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e) Tracheostomy care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f) Invasive mechanical ventilator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g) Non-invasive mechanical ventilator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h) IV medications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i) Transusions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j) Dialysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k) IV access</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

24. For the *Special Services, Treatments, and Interventions* 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):
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


———, “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, 2017a, pp. 19796–20231.


———, “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, 2017d, pp. 35270–35393.

———, “National Field Test Assessment Protocol: Admission,” November 2017e. As of January 31, 2019:

———, “National Field Test Assessment Protocol: Discharge,” November 2017f. As of January 31, 2019:

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


CMS—See Centers for Medicare & Medicaid Services.


Gage, Barbara, Laura Smith, Jessica Ross, Laurie Coots, Tracy Kline, Kate Shamsuddin, Karen Reilly, Judith Hazard Abbate, Zachariah Gage-Croll, Anne Deutsch, and Trudy Mallinson,


https://www.nap.edu/read/11623/chapter/1


Spitzer, Robert L., Kurt Kroenke, Mark Linzer, Steven R. Hahn, Janet B. W. Williams, Frank Verloin deGruy, III, David Brody, and Mark Davies, “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.


