Uniform Patient Assessment for Post-Acute Care

FINAL REPORT

January 25, 2006

Prepared by:

Division of Health Care Policy and Research
University of Colorado at Denver and Health Sciences Center
13611 East Colfax Avenue, Suite 100
Aurora, CO 80011
(303) 724-2400
Principal Investigator: Andrew Kramer, MD

Report prepared for the Centers for Medicare & Medicaid Services
and the Iowa Foundation for Medical Care

Uniform Patient Assessment for Post-Acute Care

Authors

Chapter 1
Danielle Holthaus, BS and Andrew Kramer, MD
Division of Health Care Policy and Research
University of Colorado at Denver and Health Sciences Center

Chapter 2
Barbara Gage, PhD and Jeremy Green, BA
Aging, Disability, and Long Term Care Program
RTI, International

Chapter 3
Debra Saliba, MD, MPH
UCLA Department of Medicine
VHA Geriatrics Research and Education
RAND Health

Chapter 4
Eric Coleman, MD, MPH
Divisions of Health Care Policy and Research and Geriatric Medicine
University of Colorado at Denver and Health Sciences Center

Chapter 5
Alan M. Jette, PhD, PT and Stephen M. Haley, PhD, PT
Health And Disability Research Institute
Boston University

Chapter 6
Andrew Kramer, MD
Division of Health Care Policy and Research
University of Colorado at Denver and Health Sciences Center

Editors

Andrew Kramer, MD and Danielle Holthaus, BS

January 25, 2006

Division of Health Care Policy and Research
University of Colorado at Denver and Health Sciences Center
Aurora, Colorado
303-724-2400

Uniform Patient Assessment for Post-Acute Care

Table of Contents

Page

Acknowledgements ii

Executive Summary iii

Chapter 1: Introduction, Background, and Objectives 1
Authors: Danielle Holthaus, BS and Andrew Kramer, MD

Chapter 2: The State of the Art: Current CMS PAC Instruments 15
Author: Barbara Gage, PhD and Jeremy Green, BA

Chapter 3: Review of Uniform Assessment Items and Tools for Ambulatory and Hospital Settings 39
Author: Debra Saliba, MD, MPH

Chapter 4: Hospital Discharge Assessment and Data Items that Facilitate Appropriate Placement and Efficient Care Transitions 86
Author: Eric Coleman, MD, MPH

Chapter 5: Longitudinal Outcome Monitoring across PAC Settings 100
Authors: Alan M. Jette, PhD, PT and Stephen M. Haley, PhD, PT

Chapter 6: Conclusions and Recommendations 121
Author: Andrew Kramer, MD

ACKNOWLEDGEMENTS

The authors would like to thank the following individuals for providing information and sharing perspectives relevant for the preparation of this report: Joan Buchanan; Alfred Chiplin; Richard Della Penna; Anne Deutsch; Sandra Fitzler; Meghan Gerety; Susan Hedrick; Diane Holland; David Kibbe; Carol Levine; Trudy Mallinson; Vincent Mor; Mary Naylor; Pengsheng Ni; Sue Nonemaker; David Schulke; Don Sherwood; Priscilla Shoemaker; Margaret Stineman; Steven Waldren; and Janice Zalen. We would also like to thank Karis May, Lisa Wilkenson, and Fran Weidman-Dahl for their assistance in preparing and editing this report.

Finally, we would like to thank Robert Connolly, John Burke, Judy Tobin, and Mary Pratt of the Centers for Medicare & Medicaid Services for the insights they provided throughout the development of this report.

The work described in Chapter 5 was supported, in part, by grant #AR 051870 from the National Institute of Musculoskeletal Diseases and Conditions, NIH (Dr. Jette, PI) and by an Independent Scientist Award (K02 HD45354-01) to Dr. Haley and grant R01 HD043568 (Dr. Haley, PI) from the National Institute of Child Health and Human Development (NICHD) and the Agency for Healthcare Research and Quality (AHRQ).

UNIFORM PATIENT ASSESSMENT FOR POST-ACUTE CARE
EXECUTIVE SUMMARY

Purpose

Care fragmentation, unsafe care transitions, and the inability to determine the most cost-effective settings for patients discharged to post-acute care (PAC) are all compounded by lack of a Uniform Patient Assessment. This project provides recommendations to CMS on the development of a Uniform Assessment Instrument for PAC to be completed at hospital discharge and ultimately integrated with PAC assessments. The Assessment Instrument is intended to cover the population admitted to all inpatient PAC settings (skilled nursing facilities, inpatient rehabilitation facilities, and acute long-term care hospitals), as well as residential-based PAC (home health agencies, outpatient programs). The three purposes of the PAC Assessment Instrument are: 1) placement decision-making; 2) enhancement of safety and quality of care transitions through transmission of core information to a receiving provider; and 3) provision of baseline information for longitudinal follow-up of health and function. The report was prepared by seven national PAC experts based on a review of existing instruments and literature pertinent to public and private programs, as well as discussions with other experts and CMS-recommended leaders in the health care industry. While no such review could possibly be exhaustive, every attempt was made to follow referrals and/or identify information on assessments of relevance to these purposes.

Major Findings

1. None of the three existing CMS assessment tools for PAC (MDS, OASIS, IRF-PAI) adequately covers the spectrum of patients and the necessary domains to be used across settings, and mapping across instruments is complex.
2. Past and current uniform assessment instruments (e.g., the Uniform Assessment Instrument (UNAI), Continuity of Care Record (CCR), VA Geriatrics and Extended Care (GEC) Referral Form, Health Outcomes Survey (HOS) Assessment, and others) cover some domains well, but do not yield precise measures across all patients in selected domains.
3. For the purposes of discharge planning, care transitions, and outcome assessment, a mixture of patient/proxy report measures and provider-based measures exist that could be combined from different sources to optimize data validity and minimize burden.
4. In the functional assessment domains, which are essential for uniform PAC assessment, measurement methods are in use by health systems that drastically reduce burden while improving precision of measurement across the full spectrum of impairment. These methods, termed Item Response Theory (IRT) and Computer Adaptive Technology (CAT), target the questions for an individual based on the responses to former questions so that only some items from a larger pool are answered, while scoring all persons on the same metric (see Chapter 5, page 107 for an example of IRT and CAT).

Recommendations

We recommend a two-staged development activity, lasting about one year, leading to an instrument that is ready for use in national demonstrations.

Stage 1: Instrument Development

1. Specify Domains: Thirty-one (31) domains are recommended for the three purposes of the Uniform Assessment Instrument (see Table 6.1, page 126). Although these domains were chosen based on evidence and consensus from earlier studies, a final expert panel review is recommended to assure that they fully cover the purposes of the uniform assessment instrument without excess burden.
2. Testing Functional Measurement Using IRT/CAT: The Activity Measure for Post-Acute Care (AM-PAC), developed by Boston University for functional domains and in use by Merck, HealthSouth, Kaiser Permanente of Northern California, and SeniorMetrix, could be used to demonstrate IRT/CAT for functional assessment. IRT/CAT, which would be most beneficial for measuring function, could be integrated with measures for other domains as they become available.
3. Select/Develop Measures for Each Domain: Tested and reliable measures for many domains can be adapted from existing publicly available instruments (e.g., VA GEC Referral Form, HOS) or published domain-specific measures. For some domains, additional item development will be necessary, which will require testing questions on small samples to assure both validity and reliability. Crosswalks to existing PAC instruments will be considered during measure development.
4. Automation Platforms and Transmission: We recommend a web-based approach such that the transmitting hospital can log on and conduct the assessment, which can then be accessed by the receiving provider. However, other platforms for real-time electronic data generation and transmission could be evaluated and considered in this phase.
5. Integrated Uniform Assessment Instrument: A combined uniform assessment instrument would be generated from these concurrent development and testing efforts that includes information for all domains.

Stage 2: Beta Testing

1. A sample of hospitals would be recruited and trained to complete the uniform assessment on all discharged Medicare beneficiaries. Local PAC providers would be trained to access the generated information, and development would begin on integrating the assessment information into PAC provider assessments.
2. All measures for all domains would be refined, including the metrics and item pools that are used for functional assessment using IRT/CAT.
3. Patient responses would be compared with proxy responses for patient/proxy report items.
4. Software and technology would be refined to assure that the completion, transmission, and receipt of the assessment form are as efficient as possible.
5. Care transitions would be studied for improved safety and quality.
6. Longitudinal follow-up at fixed intervals for outcome measures would be conducted to examine outcomes for different patient conditions and episodes of care.

Following beta testing, the Uniform PAC Assessment would be ready for use in national demonstration activities. By uniformly characterizing patients at hospital discharge, transmitting uniform information to receiving PAC providers, and following outcomes using the same measures over time, CMS would be able to examine quality and cost for comparable patients across PAC episodes.

Chapter 1
Introduction, Background, and Objectives

A. CONTEXT AND BACKGROUND

1. Post-Acute Care

Health care following an acute hospitalization, known as post-acute care (PAC), is provided in various inpatient settings, including skilled nursing facilities, inpatient rehabilitation facilities, and long-term care hospitals, as well residential-based settings, including home health care and outpatient rehabilitation. Skilled nursing facilities (SNFs) are nursing homes or hospital-based transitional care units that are certified by the Centers for Medicare & Medicaid Services (CMS) to provide Medicare-reimbursable skilled nursing services on an inpatient basis. Nearly 80% of Medicare residents in SNFs receive physical, occupational, and/or speech therapy,[1] and most of their nursing care is provided by certified nurse's aides. Inpatient rehabilitation facilities (IRFs) provide more intensive inpatient rehabilitation care; Medicare patients receiving inpatient rehabilitation care are required to receive a minimum of three hours of combined physical, occupational, and speech therapy per day, meet certain diagnostic criteria, and demonstrate consistent functional improvement. Inpatient rehabilitation is also characterized by much greater physician presence, particularly from rehabilitation specialists, and more care provided by licensed and registered nurses. Long-term care hospitals provide intensive care to patients who have multiple comorbidities and require inpatient hospital care over an extended period. Approximately 80% of Medicare patients in long-term care hospitals were transferred from an acute hospital.[2] Long-term care hospitals are the least frequently used PAC setting, with less 1% of beneficiaries discharged from acute hospitals using these facilities.[2] Home health care is provided to Medicare beneficiaries who are homebound (unable to leave their residences without considerable and taxing effort) and require intermittent or part-time skilled nursing care and/or therapy services. Outpatient care generally consists of follow-up physician visits and ongoing rehabilitation services or medical monitoring that are provided on an outpatient basis. In 2002, approximately one-third of Medicare beneficiaries discharged from hospitals utilized some form of PAC within one day of leaving the hospital.[3]

Within several of these PAC settings, Medicare requires that patients be evaluated using setting-specific patient assessment instruments for patient assessment, payment, and quality assurance purposes. The Minimum Data Set (MDS) is used for patients in SNFs, the Inpatient Rehabilitation Facility – Patient Assessment Instrument (IRF-PAI) is used for patients in IRFs, and the Outcome and Assessment Information Set (OASIS) is used for patients receiving home health care. (These instruments are further discussed in Section D.) Currently, no standard assessment instruments are required for patients receiving care in long-term care hospitals or outpatient care.

2. Need for Uniform Assessment Instrument

In recent years, several federal initiatives have increasingly called for development of a uniform system for patient assessment across PAC settings. The Medicare, Medicaid, and SCHIP

Benefits Improvement and Protection Act of 2000 (BIPA) mandates the Secretary of Health and Human Services to report, by January 1, 2005, on the development of health and functional assessments for various Medicare beneficiaries using PAC and other specified services. The legislation specifies that information across providers be readily comparable and that only information necessary to meet program objectives be collected. The Secretary is also required to make recommendations regarding use of patient assessment instruments for payment purposes.

In its June 2005 report to Congress, the Medicare Payment Advisory Commission (MedPAC) recommended that data elements be identified for use by CMS in establishing payments and evaluation of patient outcomes across PAC settings, asserting that the data elements "predict resource use; capture relevant clinical data; be reliable, valid, and well accepted; and minimize the burden to providers and CMS."[3, p.119]

The Institute of Medicine (IOM) has advised that "the federal government accelerate, expand, and coordinate its use of standardized performance measurement and reporting to improve health care quality," and that current performance measurement mechanisms within and across government programs be replaced by standardized measurement and reporting mechanisms.[4, p.79] The recent IOM report entitled "Performance Measurement: Acceleration Improvement," identified a lack of a "coherent, goal-oriented, consistent, and efficient system for assessing and reporting on the performance of the health care system" as one of the most significant obstacles to improving the quality of health care in the United States, calling for a concerted national effort to consolidate performance measures and reporting activities in health care.[5]

In his June 2005 testimony before the House Ways and Means Subcommittee on Health, Herb Kuhn, Director of the CMS Center for Medicare Management, stressed the importance of adopting techniques that provide more uniformity and consistency in how patients are assessed and quality is measured across post-acute settings in order to minimize care disruptions, allow sharing of patient information across settings, and avoid re-hospitalizations and other negative effects.[6] He called for collection and comparison of "consistent clinical data across various sites of service," and use of those data in an effort to build a coordinated approach to the payment and delivery of post-acute services that focus on the overall PAC episode.

Taken together, these initiatives reinforce the immediate need for a uniform system of patient assessment across sites of care – one that promotes highly consistent, coordinated, safe, and effective care for our nation's Medicare beneficiaries.

3. Report Objectives

Two central objectives guide the development of this report: a) to identify a long-term vision for uniform PAC patient assessment along with short-term, incremental steps for achieving that vision; and b) to identify a core set of potential measures or measure domains for collection across post-acute settings. These objectives are further discussed below.

a. Long-term Vision and Short-term Steps: A key objective of this report will be to present both a long-term vision for uniform PAC patient assessment and shorter-term recommendations for reaching that vision. The long-term vision will reflect the ultimate goal of the ideal strategy for uniform PAC patient assessment if we were creating the cross-site assessment without starting from the existing assessment tools. The shorter-term steps will include practical, incremental solutions toward achieving that goal over time. One short-term option, for example, while somewhat constrained, might involve testing a "cross walk" between the existing setting-specific instruments for selected domains. Other possibly intermediate solutions might involve implementation of an additional assessment – very modest in length and scope – that would contain core elements relevant to all PAC settings. A detailed discussion of specific long-term and short-term recommendations is included in Chapter 6.

b. Potential Core Measures for Collection across Settings: Rather than making recommendations for an assessment that would replace existing setting-specific instruments, the second objective of this report will be to identify a core set of potential measures or measure domains to be collected across settings – with the understanding that within each setting, providers will continue to expand on this core with unique information relevant to the patient populations served. Similarly, while the intent will not be to make recommendations for developing a comprehensive care planning tool or health record for all PAC patients, the core set of potential measures identified are intended to be relevant and useful to clinicians in providing high-quality, patient-centered care.

B. POTENTIAL PURPOSES OF A UNIFORM PAC PATIENT ASSESSMENT INSTRUMENT

Development of a uniform PAC patient assessment instrument could ultimately serve a number of purposes. The major ones include: monitoring quality and outcome across settings over time; informing decisions for placement in PAC upon hospital discharge; facilitating transitions between care settings; serving as a comprehensive care plan and clinical assessment; and serving as the basis of Medicare payment for post-acute services. Any discussion of an assessment instrument, the domains to be measured, and measurement methods should be preceded by clearly stating the instrument purposes. No instrument will be perfect for all purposes, so tradeoffs are always required when serving multiple purposes. Thus, this assessment instrument will be developed to address only certain purposes, as described below. The tradeoffs and strengths and limitations for these purposes will be discussed in Chapter 6.

1. Monitoring Quality and Outcomes across PAC Settings over Time

One potential use of a uniform assessment tool for PAC patients would be to serve as an accurate, consistent means of tracking health status and outcomes of patients as they move across post-acute settings over time. Because the current setting-specific assessment tools were created for different purposes and often assess different aspects of health and functional status (further discussed in Section D), data collected through these instruments are often not compatible – making it difficult to monitor and track patient progress (or lack thereof) from one setting to the next.

2. Informing Decisions Regarding Appropriate Placement in PAC upon Hospital Discharge

A number of studies have found that factors other than patients' clinical characteristics often play a role in determining the setting to which patients are discharged from the acute hospital.[7-10] Such factors include geographic location, availability of post-acute services, and particularly, payment related factors such as the discharging hospital's organizational structure, for-profit ownership, and ownership of PAC facilities. With different Medicare payment systems in place for each of the different post-acute settings (some of which pay per day, per discharge, or per episode), varying financial incentives exist for the provision of care across post-acute settings, particularly for hospitals that also provide post-acute, in addition to acute, health care services.

Few would argue that factors other than the patient's care needs should drive the decision regarding PAC placement upon discharge from acute care, within the constraints of provider availability. At the core of all decisions regarding subsequent care should be the best interests of the patient and his or her caregiver(s). A uniform patient assessment could serve as a patient-centered hospital discharge assessment that contains the essential elements needed for appropriate PAC placement decisions. Such a tool would establish a framework for ensuring that discharge placement decisions are based on the patient's condition, functional status, care needs, and care preferences rather than other incentives.

3. Facilitating Care Transitions across Settings

As patients are transferred from one care setting to another (or from one level of care to another within the same setting), they become particularly vulnerable to a number of potential negative outcomes, with patient safety issues and medication errors of particular concern.[11;12] The continuity of care provided within a given setting can be broken as care transitions occur, often due to problems in communicating essential information about the patient, the patient's care plan, and the care provided in the previous setting.[13] Because effective care transitions require coordination across settings as well as a reliable system for relaying key patient information, another function of a uniform assessment would be to contain a core set of key patient information that would be transferred with the patient from one setting to the next. Sharing of this core set of information would facilitate communication between care providers and help to ensure both patient safety and continuity of care.

4. Serve as a Comprehensive Care Plan and Clinical Assessment

Another potential purpose of a uniform PAC patient assessment would be to serve as a comprehensive care plan and clinical assessment for all PAC patients. While such an effort would ultimately tie in with longer-term initiatives related to electronic health records and health information technology, it would prove far too complex for the scope of this report and would therefore be better addressed within the context of other projects designed specifically for such purposes.

5. Serve as the Basis of Medicare Payment for PAC Services

The setting-specific patient assessment instruments described above serve as the basis for the Medicare prospective payment systems in their respective settings. Under prospective payment for SNFs, IRFs, and home health care, payments are established based on case mix, using a setting-specific baseline assessment and payment algorithm. For SNFs, patients are assigned to 44 groups, referred to as resource utilization groups, version III (RUG-III), which are intended to classify patients according to their needs for nursing and therapy care as indicated by the MDS. For IRFs, stays are categorized into one of 385 case-mix groups (or CMGs), which are derived based on information collected through the IRF-PAI. And for home health care, patients receiving five or more visits are allocated to one of 80 home health resource groups (HHRGs), which are determined by diagnosis, functional capacity, and service use information gathered through the OASIS.

Because the current setting-specific patient assessment instruments are used to establish Medicare payment for PAC services, the natural assumption would be that any newly developed uniform PAC patient assessment would also serve this same purpose. Such a uniform assessment would help to ensure that payment is established consistently and equitably across different sites of PAC according to the patients' clinical characteristics and health care needs, as well as services provided. This type of system would also place more of a focus on an episode of PAC, rather than the individual providers within those care episodes. While addressing issues of payment for PAC is no doubt an ultimate priority for CMS, the scope of this project does not include an exploration or direct discussion of payment-related assessment issues. However, information generated through this work will no doubt prove useful in subsequent payment-focused endeavors.

6. PAC Assessment Purposes Addressed in this Report

In summary, of the five uniform patient assessment purposes discussed above, the development of this report and the recommendations herein are guided by three. These include: 1) assessment of patients' health and functional status at the time of hospital discharge in order to support appropriate placement into post-acute settings; 2) facilitation of care transitions from the hospital to PAC and across post-acute settings; and 3) assessment of health and functional status of PAC patients across settings over time for quality monitoring.

C. EXISTING ASSESSMENT INSTRUMENTS

Each of the setting-specific patient assessment instruments introduced in Section A has unique origins in terms of conceptualization, development, and proposed uses. In addition, each was constructed using different methodology, resulting in use of different terminology, definitions, and rating scales, as well coverage of different domains and elements of care. These differences ultimately limit the tools' utility for cross-site assessment.

1. Development and History of Setting-Specific Instruments

a. MDS: As part of the Omnibus Budget Reconciliation Act of 1987 (OBRA), Congress mandated a national, uniform system for assessing all nursing facility residents by October 1990.[14] In 1988, the Health Care Financing Administration (now CMS) contracted with a consortium of researchers to develop and test a uniform resident assessment system, which culminated in creation of the MDS. The MDS was initially developed as an assessment tool to identify resident care problems that are addressed within individual residents' care plans, but over time, the use of the MDS expanded to include both payment and quality monitoring/improvement functions. As previously mentioned, information collected through MDS serves as the basis for Medicare reimbursement under the SNF prospective payment system. Data collected through the MDS are also compiled in the form of quality indicators, which are now used to monitor care quality and identify potential care problems. Quality indicator data are used by providers in their efforts to improve quality of care, by state surveyors in identifying potential problem areas for review, and by CMS for long-term quality monitoring and program planning purposes. In addition, as part of the national Nursing Home Quality Initiative launched in 2002, quality indicator information is also available to the public via the Nursing Home Care web site (www.medicare.gov/NHCompare/Home.asp).

b. IRF-PAI: The IRF-PAI consists of items from the Functional Independence Measure (FIM), which was developed as the functional assessment instrument portion of the Uniform Data System for Medical Rehabilitation (UDSmr). The concept for the UDSmr was initially developed in 1983 by a task force formed by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation. Funding for further development was provided by the National Institute of Disability and Rehabilitation Research, and the UDSmr was established by researchers at the State University of New York at Buffalo in 1984.[15] The UDSmr consists of four components: 1) a data set used to assess disability severity and medical rehabilitation outcomes; 2) computer software; 3) a data management service for subscribing facilities; and 4) a training program for users. The IRF-PAI was implemented on January 1, 2002, and now serves as the basis for Medicare reimbursement under the IRF prospective payment system. Through the UDSmr system, the IRF-PAI also serves a program evaluation function in that it allows for ongoing monitoring of patient outcomes and comparison of outcomes across facilities.

c. OASIS: Developed by the Center for Health Services Research at the University of Colorado in the late 1980s and early to mid 1990s (with funding provided by the Health Care Financing Administration and the Robert Wood Johnson Foundation) the OASIS items were designed to measure, assess, and encourage improvement in care outcomes over time using Outcome-Based Quality Improvement (OBQI) processes.[16-18] Collection and transmission of OASIS data became a requirement for all Medicare-certified home health agencies in 1999. Using OASIS data, reports presenting aggregated, risk-adjusted, descriptive, and adverse event patient outcomes are provided to all certified home health agencies; these reports permit agencies to compare their own outcomes with those of other home health agencies throughout U.S., enabling them to target their quality improvement efforts and monitor progress over time. With the implementation of the prospective payment system for home health care in October 2000, information collected via the OASIS was also used for case-mix adjustment in establishing Medicare reimbursement.

Overall, the OASIS is used for outcome monitoring, payment, and as a core but not comprehensive clinical assessment.

2. Comparison of Existing Tools

The MDS, IRF-PAI, and OASIS differ in terms of the elements they assess, their assessment periods, and their rating scales. A recent empirical comparison of the MDS, IRF-PAI, OASIS and the physical function scale of the Short-Form-36 (an assessment sometimes used with ambulatory care populations) revealed differences between and limitations within each of the instruments in terms of their content, breadth of coverage, and measurement precision.[19] An analysis by Rogers et al., found that the these tools use different terminology and definitions in describing functional ability, as well as different measurement scales for quantifying disability.[20] Similarly, a review by Bryant et al. revealed that the setting-specific instruments assess different health domains, and that items within common domains differ in perspective, qualification, source, and time periods covered.[20;21]

D. PREVIOUS EFFORTS TO DEVELOP UNIFORM PAC PATIENT ASSESSMENT INSTRUMENTS

In recent years, several initiatives have set out to develop and implement uniform systems for assessing PAC patients for a variety of purposes. Table 1.1 provides information regarding six of these initiatives, including the Continuity of Care Record (CCR),[22] the Uniform Needs Assessment Instrument (UNAI),[23] the Medicare Post-Acute Care (PAC) Quality Measurement Instruments,[24] the VA Geriatrics and Extended Care (GEC) Referral Form,[25] the Care Transitions Measure (CTM),[26] and the Personal Health Record (PHR).[27] As outlined in Table 1.1, several of these instruments – including the CCR, UNAI, VA GEC Referral, and PHR – contain key pieces of patient information needed for appropriate planning and provision of care. Others, including the CTM and Medicare PAC Quality Measurement Instruments, were designed to assess various aspects of health care quality. The CCR, UNAI, and VA GEC Referral were designed for completion by the health care provider, while the PHR, CTM, and Medicare PAC Instruments were designed to incorporate patient responses to some extent. Table 1.1 includes information regarding the intended purpose, implementation status, date developed, assessment timeframes, and domains assessed for each of these six assessment tools.

E. CHALLENGES OF UNIFORM PAC PATIENT ASSESSMENT

Development of a uniform patient assessment instrument presents many challenges, some of which include gaining provider support, minimizing burden, coordination with electronic health record initiatives, and practicalities of longitudinal follow-up. In addressing all of the following challenges as well as others, the authors' approach will be to identify short-term, practical recommendations that will bring us incrementally closer to achieving our long-term vision of uniform PAC patient assessment. This approach will provide a road map, in essence, for taking us from where we are now to an optimal uniform assessment approach across settings.

Table 1.1: Uniform PAC Patient Assessment Instruments

Continuity of Care Record (CCR)

Purpose:
- To organize and make transportable a set of basic information about a patient's health status and health care treatment that is accessible to clinicians and patients.[23]
- Intended to foster and improve continuity of patient care, reduce medical errors, improve patients' roles in managing their health, and assure at least a minimum standard of secure health information transportability.

Implementation status: Under development

When developed: Ongoing

Timeframe for use: At conclusion of a health care encounter

Domains Covered:
- Patient information
- Provider information
- Insurance information
- Allergies/alerts
- Advance Directives
- Diagnosis/problems/conditions
- Medications
- Immunizations
- Vital signs
- Procedures
- Lab results
- Encounters
- Social history
- Family history

Uniform Needs Assessment Instrument (UNAI)

Purpose:
- To determine an individual's needs for continuing care.[28]
- *Not* developed to replace a comprehensive geriatric assessment, to serve as a care plan, or to be the primary means for determining eligibility for Medicare covered services.

Implementation Status:
- Not implemented on a national scale.
- Field tested in 1999.
- Under consideration by CMS for use (in possibly modified form) as part of discharge planning process to determine appropriate post-acute setting, 2005.

When Developed:
- Originally developed in 1992
- Refined in April 1999

Timeframe for use: Prior to discharge

Domains Covered:
- Identification information
- Health status and medical complexity
- Cognitive status
- Sensory status and communication
- Mental health and behavioral factors
- Physical functioning
- Continence
- Resources and goals for discharge
- Nursing and other care requirements at discharge

Medicare Post-Acute Care (PAC) Quality Measurement Instruments

Purpose:
To measure the quality of PAC for patients with stroke, congestive heart failure, pneumonia, and back and neck conditions (as tracer conditions).

Implementation status:
- Small-scale feasibility tests conducted as part of developmental process.
- Large-scale implementation of stroke instrument through multi-state study of Medicare PAC for stroke patients, 2002-2005.

When developed: 2001

Timeframe for use:
- PAC admission
- 90 days following PAC admission

Domains Covered:
- Physical function outcomes
- Mental health outcomes
- Quality of life outcomes
- Utilization outcomes
- Physiology outcomes
- Satisfaction outcomes
- Process of care

Table 1.1: Uniform PAC Patient Assessment Instruments (continued)

VA Geriatrics and Extended Care (GEC) Referral Form

Purpose:
- To identify long-term care needs and assure that each patient receives necessary services at the least restrictive level of care.[29]
- To be used as a screening tool, not a full assessment.[30]

Implementation status: National implementation planned for February or March 2006.

When developed: 2000

Timeframe for use: Prior to placement in long-term care

Domains Covered:
- Source of referral
- Living situation
- Primary caregiver information
- Language
- Homebound status
- Instrumental activities of daily living
- Services in the home
- Skilled care
- Basic activities of daily living
- Continence
- Skin
- Patient behaviors and symptoms
- Cognitive status
- Prognosis
- Weight bearing
- Diet
- Equipment/supplies needed
- Goals of care
- Program referring to; estimated duration of services

Care Transitions Measure (CTM)

Purpose:
- To assess the quality of care transitions for the purpose of performance measurement and subsequent public reporting.[26]
- For use in promoting quality improvement in transitional care.

Implementation status:
- First used in a quality improvement project in 2000.
- Currently being used in at least 6 care transition quality improvement projects.

When developed: 1999

Timeframe for use: Following a health care transition (14-28 days post-transition)

Domains Covered:
- Understanding one's self-care role in the post-hospital setting
- Medication management
- Having one's preferences incorporated into the care plan

Personal Health Record (PHR)

Purpose:
To encourage patient "ownership" of a core set of information important for facilitating cross-site communication and continuity of care across providers and settings.

Implementation status: Openly accessible for patients to download or for practitioners to download to share with patients.

When developed: 2001

Timeframe for use: Upon discharge from hospital and updated by the patient and family caregiver on an ongoing basis

Domains Covered:
- Patient information (including advance directives and PCP info.)
- Hospitalization information
- Caregiver information
- Medical history
- Medication record
- Allergies
- Tasks to complete before discharge

1. Gaining Provider Support

In order for any uniform assessment instrument to be effectively implemented, it will be imperative that both acute and post-acute providers support the use of the instrument. Critical steps toward achieving this support will be to ensure that the data elements incorporated are useful to providers in carrying out their day-to-day patient care responsibilities (i.e., in terms of care planning, improving quality of care, etc.) and that the burden of data collection is minimized. Financial rewards based on completion of the instrument (i.e., through a pay-for-performance initiative) would also provide an additional incentive for provider compliance. Engaging provider representatives in the development process always enhances buy-in, but must be managed so that the process does not become prolonged.

2. Minimizing Burden

A key factor in the success of uniform assessment instrument will be to keep burden of data collection (for both patients and providers) to a minimum. By including only a core set of essential measures to be collected across post-acute settings – which can be supplemented by relevant setting-specific data elements as deemed necessary – the length of the instrument and the time required for completion can be minimized. However, we are trying to address three different purposes with the instrument, which will require tradeoffs to keep burden to a minimum.

3. Coordination with Electronic Health Record Initiatives

With increasing importance being placed on the use of health information technology standards and electronic health records to improve the health care quality, coordination, and efficiency of care, this movement must be considered in the development and implementation of a uniform PAC assessment instrument. While the focus of this report will not be to address the specifics of electronic data collection, entry, and transmission, the authors recognize the importance of formulating recommendations that are complementary to and consistent with the aims of health information technology and automation.

4. Practicalities of longitudinal follow-up

In devising a uniform tool for quality and outcome monitoring across post-acute settings over time, key decisions will have to be made in terms of appropriate intervals for assessment, appropriate respondents (i.e., provider, patient, proxy, or a combination of the three), and mode of data collection. Such decisions are complex and will require thorough consideration of many factors, including how the results will be reported and when, who will be responsible for collecting and assuring quality of the data, and psychometric and quality measurement issues.

F. REPORT OVERVIEW

The remainder of this report is divided into five main chapters, brief overviews of which are provided below.

The remainder of this report is divided into five main chapters, brief overviews of which are provided below.

Chapter 2 will include an in-depth review of the existing CMS instruments (MDS, OASIS, IRF-PAI), the extent to which they equate to one another, and the strengths and limitations of each. Chapter 3 will discuss key domains for use in assessing health and functional status to support appropriate PAC placement, facilitate care transitions, and monitor quality over time. Recommendations of potentially useful measures for inclusion in a uniform patient assessment instrument will be made. Chapter 4 will build upon Chapters 2 and 3 to recommend a uniform assessment approach for use in making appropriate PAC placement decisions upon hospital discharge and in facilitating care transitions for patients from the hospital to PAC and across post-acute settings. Chapter 5 will make recommendations for a uniform longitudinal assessment to monitor health and functional status of PAC patients across settings over time. Finally, Chapter 6 will present a distillation of findings and a set of conclusions and recommendations for the development of a uniform PAC assessment instrument.

G. REFERENCES

1. DataPRO Team. Skilled nursing facilities prospective payment system quality medical review Data Analysis Pro: Final Report for CMS Contract 500-99-CO01, Attachment J-19. 2002.

2. Medicare Payment Advisory Commission. Report to Congress: variation and innovation in Medicare, Chapter 5 - monitoring post-acute care. 2003.

3. Medicare Payment Advisory Commission. Report to the Congress: issues in a modernized Medicare program. Chapter 5: Payment for post-acute care. 2005.

4. Institute of Medicine. Leadership by example: coordinating government roles in improving health care quality. 2005. National Academies Press.

5. Institute of Medicine. Performance measurement: accelerating improvement. 2006. Washington, DC, National Academies Press.

6 Kuhn H. Testimony of Herb Kuhn, Director, Center for Medicare Management, Centers for Medicare & Medicaid Services. Hearing on Post-Acute Care. 6-16-2005. House Committee on Ways and Means, Subcommittee on Health.

7. Bronskill SE, Normand ST, McNeil BJ. Post-acute service use following acute myocardial infarction in the elderly. Health Care Fin Rev 2002; 24(2):77-93.

8. Kane RL, Wen-Chieh L, Blewett LA. Geographic variation in the use of post-acute care. Health Serv Res 2002; 37(3):667-682.

9. Gage B. Impact of the BBA on post-acute utilization. Health Care Fin Rev 1999; 20(4):103.

10. Lee AJ, Huber JH, Stason WB. Poststroke rehabilitation in older Americans. The Medicare experience. Med Care 1996; 34(8):811-825.

11. Coleman EA, Berenson RA. Lost in transition: challenges and opportunities for improving the quality of transitional care. Ann Intern Med 2004; 141(7):533-536.

12. Moore C, Wisnivesky J, Williams S, McGinn T. Medical errors related to discontinuity of care from an inpatient to an outpatient setting. J Gen Intern Med 2003; 18:646-651.

13. Coleman EA. Falling through the cracks: challenges and opportunities for improving transitional care for persons with continuous complex care needs. J Am Geriatr Soc 2003; 51(4):549-555.

14. Morris JN, Hawes C, Fries BE, Phillips CD, Mor V, Katz S et al. Designing the national resident assessment instrument for nursing homes. Gerontologist 1990; 30(3):293-307.

15. Granger CV, Brownscheidle CM. Outcome measurment in medical rehabilitation. Int J Technol Assess Health Care 1995; 11(2):262-268.

16. Shaughnessy PW, Crisler KS, Schlenker RE, Arnold AG. Outcomes across the care continuum: home health care. Med Care 1997; 35(11):NS115-NS123.

17. Shaughnessy PW, Crisler KS, Schlenker RE. Outcome-based quality improvement in home health care: the OASIS indicators. Qual Management Health Care 1998; 7(1):58-67.

18. Shaughnessy PW, Crisler KS, Schlenker RE, Arnold AG, Kramer AM, Powell MC et al. Measuring and assuring the quality of home health care. Health Care Fin Rev 1994; 16(1):35-67.

19. Jette AM, Haley SM, Ni P. Comparison of functional status tools used in post-acute care. Health Care Fin Rev 2003; 24(3):13-24.

20. Rogers JC, Green Gwinn SM, Holm MB. Comparing activities of daily living assessment instruments: FIM, MDS, OASIS, MDS-PAC. Phys and Occup Ther Geriatr 2001; 18(3):1-25.

21. Bryant LL, Floersch N, Richard AA, Schlenker RE. Measuring healthcare outcomes to improve quality of care across post-acute care provider settings. J Nurse Care Qual 2004; 19(4):368-376.

22. Secretary's Advisory Panel on the Development of Uniform Needs Assessment Instrument(s) USDoHaHS. Uniform Needs Assessment Instrument (UNAI). 1993.

23. ASTM International. Work Item Summary (WK 4363): Standard specification for the Continuity of Care Record (CCR). 2005.

24. Johnson M, Holthaus D, Harvell J, Coleman E, Eilertsen T, Kramer AM. Medicare post-acute care quality measurement: final report. 2001. Denver, Colorado, Center on Aging Research Section, University of Colorado Health Sciences Center.

25. Office of Geriatrics and Extended Care DoVA. VA Geriatrics and Extended Care Referral Form. 2005.

26. Coleman EA, Mahoney E, Parry C. Assessing the quality of preparation for posthospital care from the patient's perspective: the care transitions measure. Med Care 2005; 43(3):246-255.

27. Coleman EA, Smith JD, Frank JC, Min S, Parry C, Kramer AM. Preparing patients and caregivers to participate in care delivered across settings: the care transitions intervention. J Am Geriatr Soc 2004; 52(11):1817-1825.

28. CMS Office of Clinical Standards and Quality Quality Management and Health Assessment Group. Uniform Needs Assessment Instrument Overview. 8-17-2005.

29. Office of Geriatrics and Extended Care, Department of Veterans Affairs. GEC Referral Form Manual. 1-30-2005.

30. Office of Geriatrics and Extended Care, Department of Veterans Affairs. GEC Referral Form Frequently Asked Questions. 2005.

Chapter 2
The State of the Art: Current CMS PAC Instruments

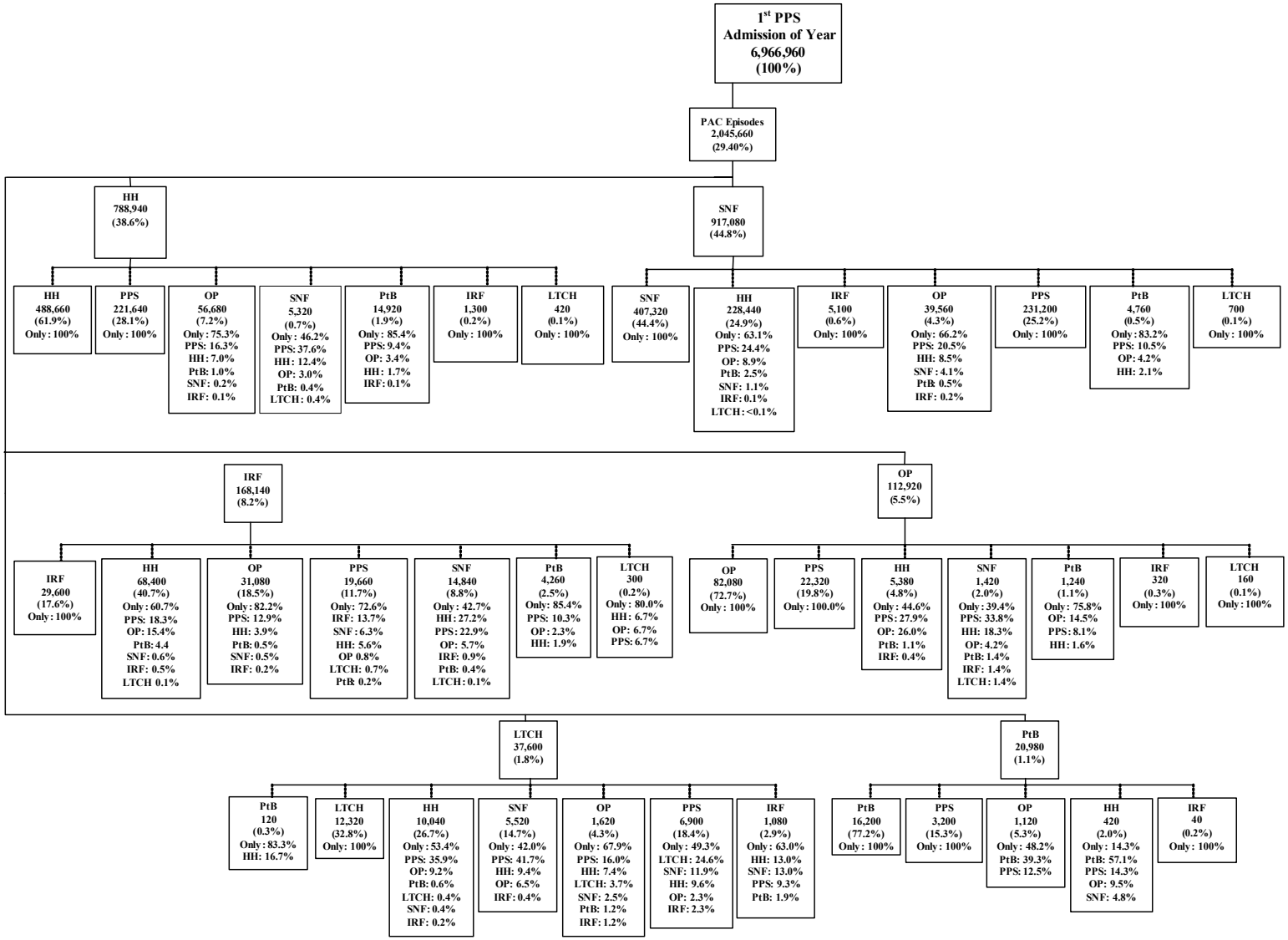
A. INTRODUCTION

Medicare provides insurance coverage for many types of care, which when used in succession constitute episodes of care. About 20 percent of all beneficiaries are admitted to a general acute hospital during a year.[1] In addition, Medicare covers other types of hospitalizations which are used less frequently by the general populations, but for hospital discharges with certain conditions, a high proportion may be transferred from the general acute or directly admitted to inpatient rehabilitation facilities (IRFs), psychiatric hospitals, and long- term care hospitals. Many hospitalized beneficiaries are transferred to a skilled nursing facility (SNF). More importantly, all SNF admissions are transferred *in from* one of these hospital beds. About 39 percent of the hospital discharges will go on to use home health agency services during a year. These services may be in addition to, or in place of outpatient therapy services. The outpatient services may be accessed through a hospital outpatient department, rehabilitation agency, comprehensive outpatient rehabilitation facility, or independent therapist's office.[2] These sequences make up typical PAC episodes (Figure 2.1).

While beneficiaries may transfer across this continuum or enter PAC at different points, these providers are generally expected to be delivering different levels of care. General acute hospitals are seen as sites that treat acute illnesses or acute exacerbations of chronic illnesses. They diagnose, treat, and discharge patients typically within four to six days.[3] Long-term care hospitals are also acute hospitals but their patients typically have longer lengths of stay, averaging admissions of about 25 days. They specialize in patients with respiratory conditions, rehabilitation needs, or psychiatric illnesses.[4,5] IRFs, on the other hand, specialize in only acute patients who need inpatient-level, intensive rehabilitation therapy. Their admissions are comprised primarily of orthopedic, neurological, and other specialized cases requiring physical therapy, occupational therapy, or nursing (both rehabilitation and medical).[6] SNFs admit patients who may need less intensive rehabilitation therapy following IRF service use, or who may be too ill to be admitted to an IRF and/or need nursing services and conditioning. Outpatient therapy settings and home health agencies provide these same therapy services to patients who do not require inpatient care; those who are home-bound may also be accessing nursing, home health aide, and social worker services in their home.

Many of these providers have the same types of staff but provide different levels of services.[7] Each have physical therapists, speech and language pathologists, occupational therapists, nurses, and physicians but the intensity of the service provided varies by setting.[8] Because of these overlaps, the Medicare Payment Assessment Commission (MedPAC) and its predecessor, the Prospective Payment Assessment Commission (ProPAC) have been calling for consistent measurement across settings since at least the mid-1990s.

Figure 2.1: Post Acute Care Transitions, 2002



NOTE: Sample data were inflated to estimate total Medicare service utilization
SOURCE: Gage, Bartosch, Leung, Pilkauskas, Hoover, and Green, 2005b.

But much of the early research to develop PAC case mix measurement systems has been done by separate research groups, each with expertise in their respective PAC service area. Hence, as the new PAC prospective payment systems (PPS) went into law in 1997, CMS used the case mix systems, payment methods, and measurement tools that had been developed for each service over the prior decades.

Today, the general acute, long-term care hospital (LTCH), IRF, SNF, home health (HH), outpatient, and psychiatric payment systems each have their own, distinctive case mix systems. The most closely aligned are the general and long-term acute hospitals that both use the same diagnosis-related groups (DRGs), although the payment weights differ for the two payment systems. The psychiatric hospital PPS builds on this system but adjusts it with measures appropriate for those populations. The IRF, SNF, and HH systems are independent, despite the overlap in types of services provided, and none are based on the DRG system. Each was developed by experts in their respective service areas.

This chapter compares the patient assessment tools currently used in the Medicare program to measure case mix differences within each PAC setting and contrasts them across settings. These tools are important for monitoring beneficiary access, quality of care, and appropriateness of admissions in each setting. Ideally, if the measurement systems were consistent across settings, these types of tools could be used to monitor patient transfers and outcomes across settings. However, none of these tools were designed to measure patients across settings.

This chapter will compare the tools currently in place, identify similarities and differences in measurement and application, and discuss the research that has been conducted or is needed before a uniform tool is available for PAC assessments.

B. OVERVIEW OF THE TOOLS

The BBA and subsequent legislation established several new PPSs that use patient assessment data to adjust base payment rates for each respective provider. However, each system uses a different patient assessment tool:

- Skilled nursing facilities use the MDS or minimum data set for skilled nursing facilities;
- Home health agencies use the OASIS or Outcomes and Assessment Information Set tool;
- Inpatient Rehabilitation Facilities use the IRF-PAI or IRF Patient Assessment Instrument.

Long-term care hospitals do not have an assessment tool but the Centers for Medicare and Medicaid Services (CMS) are currently conducting a study to develop criteria to distinguish these hospitals from other types of acute inpatient hospitals.[7] This study is examining the tools used in the private sector to determine "level of care" definitions. These levels are used by the quality improvement organizations (QIOs) and some private insurers as guidelines to authorize services in different settings. They are based on medical severity, functional impairment levels, and treatment needs. Outpatient therapy providers do not submit patient assessment information but the SF-36 is sometimes used to measure function in this population.

1. History and Development

As noted in Chapter 1, the three tools (MDS, OASIS, and IRF-PAI) were developed by different parties for different purposes. While two of the tools, the MDS and IRF-PAI, were developed for use with inpatient populations, these populations differ dramatically. The MDS was originally developed for a long-term care population with high cognitive and functional impairments while the IRF tool was developed for use with an acute, inpatient rehabilitation population. The latter had to be strong enough to undergo at least three hours therapy across five days/week. The third tool, OASIS was developed for use with homebound, post-acute populations. Both the IRF and HH populations are relatively healthy compared to an institutionalized long-term care resident. While the MDS has evolved over time, the original form was emphasizing factors related to frailty while the other two forms have always been applied to more acute populations. The MDS now applies to both LTC and post-acute populations in nursing facilities. Each of these tools have different histories.

a. The MDS: The MDS grew out of the long-term care field as a care planning and quality monitoring tool for long-term care residents in nursing facilities. The Omnibus Budget Reconciliation Act of 1987 (OBRA 1987) directed CMS (then HCFA) to develop a care planning tool for long-term care facilities participating in the Medicaid program. The law identified specific components to be included in such a tool. This tool incorporated recommendations from hundreds of experts in geriatrics, psychiatry, therapy, social work and residents rights advocates (Mor, 2005).[9] CMS funded a consortium of researchers and practitioners to develop what became known as the MDS for nursing homes.(1) The Katz ADL scales were used as the basis of this instrument (Katz, 1983).[10] OBRA 1987 further mandated that a data collection system be established to monitor quality of care in long-term care facilities, particularly in the Medicaid program.

As policymakers began discussing a PPS for SNFs, the MDS was turned into the existing instrument used in facilities that often have both long-term care and skilled nursing units.(2) [11] Variations of the MDS were tested with VA and other populations and eventually led to the development of companion tools for home health (MDS-HC) and other services. Today, these have evolved into a suite of tools developed by the Resident Assessment Instrument (RAI) consortium to measure functional improvement potential, institutional risk, communication disorders, cognition, depression and anxiety, social function, cardio-respiratory, dehydration, falls, nutrition, pressure ulcers, palliative care, reduction of formal services, and urinary incontinence.[12] In general, MDS items focus on health conditions, cognitive impairments, other complicating conditions, such as communication, psychiatric, and pain issues as well as necessary nutritional approaches other than self-feeding, skin conditions, and special treatment needs, such as chemotherapy, dialysis, ostomies, tracheostomies, transfusions, ventilator management, Alzheimer's units, podiatry, or hospice care; in other words, frailty factors that are more common in the long-term care population. The MDS documents other frailty factors such as whether the patient is comatose, delirious or has behavioral symptoms, such as wandering,

(1) This consortium was lead by Catherine Hawes of RTI, International and included Brant Fries of the University of Michigan, John Morris of the Hebrew Home for the Aged, and Vincent Mor of Brown University.

(2) In 1992, the Prospective Payment Assessment Commission reviewed the range of tools used in nursing facilities. (See ProPAC, 1992).

agitation, or other factors that would require restraints or closer monitoring. CMS is currently considering revisions to the MDS and where applicable, those modifications will be discussed here.

b. The OASIS: As noted earlier, the OASIS was originally developed by the University of Colorado as an outcome assessment tool for home health care. Most of the items were developed under research co-funded by CMS and the Robert Wood Johnson Foundation. The tool, which evolved over the last decade, now includes additional items that are useful for clinical assessment and care planning in addition to outcomes monitoring, but it is not intended as a comprehensive assessment.

The target population is in a home environment where issues such as personal assistance and environmental safety are as important as individual health status and functional status. This tool collects data on living arrangements, medical conditions, neurological, emotional, behavioral conditions, ADLs, IADLs, and self-management issues, such as patients' and caregivers' abilities to manage medications and equipment, such as oxygen, IVs, and tubes. In addition, data are collected on prior hospitalization and emergent care needs.

Like the MDS, assessments are primarily conducted by nursing staff although therapists also contribute where appropriate. The OASIS can be used to help the nurse assessor develop the patient care plan, including requests for assessments by other professionals, such as occupational therapists to assess the home environment, physical therapists to assess the level of therapy needed (if any), social workers to assess the family and other types of support available to the individual, and aides, including the types of assistance needed.[3] This same tool is used for planning services in both the Medicare and Medicaid programs despite differences in the extent of insurance coverage for unskilled assistance, such as home health aides.

c. The IRF-PAI: Medicare's IRF-PAI tool grew out of an effort to improve measurement of functional impairment for inpatient rehabilitation patients with various diagnoses, but all of whom needed physical medicine and rehabilitation. A task force was formed in the early 1980s, which reviewed existing scales and functional assessment instruments. Under funding from the National Institutes of Disability and Rehabilitation Research (NIDRR) coupled with support from the Robert Wood Johnson Foundation (RWJ), an effort was made to select the most common and useful items for rehabilitation clinicians to assess severity of disability in a uniform and reliable manner. The final result, the Functional Independence Measures (FIM)^{TM} was based on the Barthel Index[13] modified to include communication and cognition items. The scales were modified from three levels to seven levels, and the weights were removed.[14]

The FIM^{TM} tool is an additive scale comprised of 13 physical motor skill items and 5 cognitive or communication items, each rated on a seven-point scale measuring degree of dependence and frequency of need for assistance. The IRF-PAI supplements these items with diagnostic information, personal support information, and certain quality indicators. The focus is on measuring impairments in a group that is strong enough (i.e., not frail) to participate in intensive

[3] Medicare only covers aide services if intermittent or part time skilled services are needed (nursing or physical therapy or on-going occupational therapy). Medicaid, on the other hand, provides home health aide services in many states.

therapy nearly every day. CMS has recently completed a study to modify the items to better measure quality in the IRF setting. Where applicable, those modifications will also be introduced in the following discussions.

These assessment tools must be completed on every Medicare admission in the respective setting. The MDS is completed on all Medicaid and Medicare covered nursing facility admissions, the IRF-PAI is completed on all Medicare covered inpatient rehabilitation facility admissions, and the OASIS is completed on all Medicaid and Medicare covered home health patients.

2. General Measurement Issues

While the three tools were designed to assess different populations, they all collect the same *types* of information to varying degrees. These can be grouped in five basic categories:

- Administrative information
  - Demographics
  - Insurance coverage
- Social support/residency prior to admission
  - Type of Residence
  - Level and Type of Assistance Received
- Medical diagnosis and conditions
  - Active medical conditions
  - Complicating conditions, such as skin, respiratory, pain issues
  - Special equipment needs (tubes, oxygen, dialysis, etc.)
- Functional limitations
  - Activities of Daily Living
  - Instrumental Activities of Daily Living
  - Physical mobility
  - Cognitive impairments

a. Tool Composition: While these five categories are shared by the assessment tools, the domains, actual items, item definitions, scoring methods, and metrics differ across tools.[15] Some differences exist because of the emphasis of a respective service. For example, the IRF-PAI is the shortest form and emphasizes the medical and functional needs of the patient. Its functional measures (FIM scores) are based on a modified Barthel Index which builds on the ADL concepts but measures what a patient does and their need for supervision, not a measure of what the patient is capable of doing and expands the mobility measures to better capture small increments of change in the rehabilitation population. The OASIS, which is used in the home environment, is a longer form because it includes the health and functional domains but also adds items on need for (Instrumental Activities of Daily Living) and level of assistance available in the home.

In addition, the OASIS collects information on the medical conditions for which the patient was recently hospitalized since these may not be the reason for the HH admission but they could affect the HH treatments. The MDS is the longest form and it collects a limited number of medical and functional impairment measures but a large number of items on special equipment, neurobehavioral issues, such as cognitive impairments, memory recall, and propensity to wander. Thus, even when the domains of health and function are consistent across tools, many of the items used to measure them differ. Each of the items will be discussed in more detail below as they affect the viability of a uniform PAC assessment system based on today's tools.

b. Assessment Timing: In addition to specific items varying across tools, the three systems also differ in terms of when a patient is assessed. Ideally, this information would be collected at admission and discharge so one could measure the severity of illness or impairment levels and use them to determine appropriate admissions, appropriate discharges, and to measure the impact of each treatment.

However, each tool collects data on a different schedule (Table 2.1). While all three tools collect data "at admission," only the OASIS assessor records on the day of admission. In IRFs, the admission period may vary across three days and in SNFs, across eight days. The SNF admission has interim assessments at 14 days and 30 days and every 30 days after that until coverage ends at day 100. The OASIS has a follow up assessment every 60 days and both OASIS and IRF-PAI include a discharge assessment, but one is not required in the SNF.

Table 2.1: Frequency, time period covered, and measurement scales differ across post-acute patient assessment tools required by Medicare

Dimension	Inpatient rehabilitation facilities	Skilled nursing facilities	Home health agencies	Long-term care hospitals
Tool	IRF-PAI	MDS	OASIS	None
Frequency of assessments	At admission and discharge	Initial (day 1-8); day 14; day 30; and every 30 days, up to day 100	Initial at admission; every 60 days thereafter; and at discharge	
Time period covered	Lowest level within first/last 3 days (for admission and at discharge)	Generally 7-day look-back	Status of patient on day of assessment	
Method of assessment	Directed observation preferred but can be combined with reported performance	Information gathered for multiple caregivers' descriptions and documentation. Direct observation not required.	Direct observation preferred, but also often used interviews with patient in-home caregiver	
Minutes to complete	25 minutes	90 minutes	90 minutes	

NOTES:
MDS (Minimum Data Set), OASIS (Outcome and Assessment Information Set), IRF-PAI (Inpatient Rehabilitation Facility-Patient Assessment Instrument).

SOURCE: IRF-PAI Training Manual.[16] and MedPAC Report to the Congress[17]

The tools also vary in terms of whether the assessment is judging health and impairment levels on the day of assessment or for some prior period, as in the case of the IRF-PAI and MDS assessments. This can lead to more subjective recordings than direct observation at the time of assessment. According to past studies, the three tools also vary in how long it takes to complete them.[17]

3. Item Comparisons

While many of the domains are similar across the tools, the actual items used to measure the domains differ substantially. Each of these areas are compared below (Table 2.2).

Table 2.2: Overview of Selected Domains and Items in the MDS, IRF-PAI, and OASIS Tools

	SNF	IRF-PAI	OASIS
I. Administrative Information			
Age	X	X	X
Race	X	X	X
Social Security No.	X	X	X
Medicaid No.	X	X	X
Marital status	X	X	X
II. Social Support/Residency			
A. Living Situation			
Prior Residence	AB.5.All settings in 5 years prior to NF residence	16: Prehospital Living Setting (same codes as "admitted from"	M0300: current residence
Setting Discharged to		44a. Discharged to	
Prior Lives With	AB.3. Lives alone (Y/N/in facility)	17. Pre-hospital living with (alone/family or relatives/friends/attendant/ other)	N.A.
B. Assistance Available			
Lives with after discharge	AB.3. Lives alone (Y/N/in facility)	45: Alone, relations, friends, paid	MO340: Alone, spouse, other family, friend, paid
NonHC Assisting person	----------	----------	M0350
Type of Primary Caregiver	----------	----------	M0360
Frequency of PC Assistance	----------	----------	M0370
Type of Assistance Used	----------	----------	M0380: ADL, IADL, Environmental, Psychosocial, Advocates

Table 2.2: Overview of Selected Domains and Items in the MDS, IRF-PAI, and OASIS Tools (continued)

	SNF	IRF-PAI	OASIS
III. Medical Conditions			
A. Diagnosis			
Primary Diagnosis	Section I.1. Check all listed diagnosis that affect ADL, cognitive, behavior, medical treatment, need for monitoring, or risk of death (set of conditions)	22: Etiologic Diagnosis (ICD- 9 for admission)	M0230: Primary Diagnosis ICD-9 code
Prior inpatient-last 14 days	(Transferring information)	----------	M0190:ICD-9
Treatment change last 14 days	----------	----------	M0210:ICD-9
Severity Rating	Section J.5.: Stability of general condition affecting a) behavior, b) exacerbated acute/chronic problem, c) end stage disease, d) none affected	None	M0230/240: 4 point item measuring symptom control for each primary and other ICD-9
B. Comorbidities/ Complications	Section I. Check all on the list that apply, no primary diagnosis specified	24: Up to 10 ICD-9 being treated and 47: up to 6 ICD-9s that began with IRF admission	M0240: other diagnosis
C. Specific Compli-cating Conditions			
Urinary	Several sections, including G.i. toilet use; section H: bowel and bladder issues (5 point scale of continent, usually (LT 1/wk), occasionally incontinent (bladder 2+/wk, bowel (1/wk), frequently incontinent (daily bladder with some control, bowel2-3/wk), incontinent (multiple daily, almost all) and appliances or programs used	Items 29-39 (FIMtm modifiers and Sphincter control items (7 point scale measuring complete independence, modified (uses device), supervised 100%), minimal assist (75%+, moderate assist (50%+, maximal assist (25%+, total assist).	M0220: lists 6 conditions in past 14 days, including: incontinence, indwelling catheter M0510-M0550 asks if in the last 14 days, UTI treated, incontinence or catheter used, frequency/ timing of urinary and bowel incontinence, and whether patient had an ostomy.

Table 2.2: Overview of Selected Domains and Items in the MDS, IRF-PAI, and OASIS Tools (continued)

	SNF	IRF-PAI	OASIS
Pain	J.2. Frequency daily or less, Intensity of mild, moderate, excruciating	51. Rate 1-10 over 3 days	M0420: Frequency, M0430: Intractable
Cognitive impairment	B4: 3 point decision impairment scale E4: Symptoms over last 7 days B2-3: Recall last 7 days	See below and FIMs items P-R.	Impaired decisions Disruptive/inappropriate behavior Memory Loss needing supervision
Delirium	B5: Delirium in last 7 days (6 items/3 point scale)	26. At admission	----------
Comatose	B.1.(Y/N)	25. At admission	----------
Dehydrated	J.1c.: As part of problem conditions	28. At admission/ discharge	M0840:As 1 of 9 reasons for emergent care
Accidents/falls	J.4.:Fall/fx in past 180 days	53/54: Balance problem or number of falls during IRF stay	
Respiratory Status	J.1.l (shortness of breath)		
Shortness of breath: exertion		48 (N/Y, Adm/Dc)	M0490
Shortness of breath: at rest	J1.b (inability to lie flat due to shortness of breath)	49 (N/Y, Adm/Dc)	----------
Cough/secretions		50 (N/Y, Adm/Dc)	----------
Respiratory treatments	Section P.1. (oxygen, suctioning, tracheostomy, ventilator/respirator)	----------	M0500, MO44OS MO440S, 450S, 460S, 470S, 480S
Pressure Ulcers	Section M1-6 (#/stage, stage/type, history, skin treatments, foot problems)	52A: Highest stage, 52B: Number 53: PUSH Tool	M0440-M0488 (total number/stage, and #/stage of worst by type

Table 2.2: Overview of Selected Domains and Items in the MDS, IRF-PAI, and OASIS Tools (continued)

	SNF	IRF-PAI	OASIS
D. Sensory			
Swallowing status	K1 (swallowing/chewing) K5 (nutritional approach)	27: Regular/ modified/tube/PTN feeding	M0710 (Feeding/eating abilities, prior/current)
Vision	D1 (ability to see – 5 level scale)	----------	M0390
Hearing		----------	M0400
Speech			M0410

NOTE**:** Initial numbers identify item number in the respective tools.

a. Administrative Information: Each tool collects basic demographic information on the patient including age, gender, social security and insurance numbers. These data are fairly standard across all tools and could be used to merge datasets.

b. Social Support Systems/ Residence Prior to Admission: The types of social support questions vary across the three tools. The two tools used with the healthiest population, the IRF-PAI and OASIS, both identify where the patient lived prior to admission, who they lived with, and in the OASIS, the degree of help these people provided to the patient prior to admission. The IRF-PAI also asks where a patient will be discharged and with whom they will live after discharge. The latter does not specify whether the co-residents will provide assistance or the frequency of assistance available, which is the information the therapy staff needs to understand the level of achievable independence.[4] In contrast, the MDS collects information on the resident's residential history over the five years prior to admission, whether they lived alone (Y/N/Other facility), and upon discharge, the type of setting they will be entering. As reflected in the variation of the current tools, certain settings require additional information that may not be relevant to all settings, such as the IADLs in the OASIS. However, to the extent they are collecting the same types of information, the current tools do not currently use the same items.

Despite the similarities in the domains in the IRF-PAI and OASIS, they currently ask very different questions. The IRF-PAI collects information on the patient's pre-hospital living setting. Included in the response groups are both community-based residential options and responses answering where the patient was admitted from, including three types of hospitalizations (IRF, LTCH, and general acute).[5] The OASIS questions target mostly community-based options to determine whether the person lives alone or with others, and if the latter, the relationships they have. Both forms ask who the patient lives with but the responses are coded differently, with the

[4] RTI proposed changing this item to be more consistent with the OASIS item in its IRF-PAI modification recommendations to CMS.

[5] RTI also recommended to CMS that this item be replaced with the OASIS items.

IRF-PAI only distinguishing between being alone, with any family, with friends or with a hired attendant. The OASIS categories are similar conceptually but offer different choices, which identify whether the family member is a spouse. These two items could conceivably be crosswalked, or alternatively, if the IRF-PAI QI recommendations go into effect, both forms will use the same codes.

More importantly, OASIS asks about the type of assistance given by the primary caregiver and by whom. Again, it was recommended that these items be added to the IRF-PAI since they describe the type and frequency of resources that will be available if they are discharged home. The extent of available help is significant in predicting the probability of discharge home from an IRF.[6] In contrast, the IRF-PAI contains discharge items that collect the same information as the admission set (i.e., discharged where and to live with whom) but this does not describe the degree to which their co-residents will be willing to help them.[6] The MDS collects information on where the patient is admitted from and their discharge status. While the admission and discharge groups are consistent within the MDS, they offer more choices than the IRF-PAI or OASIS in terms of subsequent health care settings. This information, if reliable and consistent in the future, could be useful for tracking utilization patterns and care transitions across setting.

c. Medical Conditions: Each tool collects information on the patient/resident's medical conditions. They differ in whether they specify the primary condition for which the patient is being treated and the degree to which they use open-ended items. Close-ended items increase the ability to compare across patients and tools but they also include some assumptions regarding the most common conditions.

d. Active Medical Conditions: The IRF-PAI and OASIS use open-ended items requesting ICD-9 codes to identify the medical conditions that need to be treated while the patient is receiving that service. In addition, the OASIS also asks about recent past diagnosis. First, the OASIS asks for a list of ICD-9 codes for conditions treated in an inpatient facility during the last 14 days (M0190), the codes requiring a change in HH treatment regimen (M0210), and the codes for which the patient is being treated in home health (M0230/0240). This last one has a four-point severity index applied to it which designates whether the symptoms are controlled and to what degree they need adjustment or have had related rehospitalizations because of it.

In contrast, the IRF-PAI collects only the ICD-9 codes related to the admission (22) and up to 10 comorbid conditions (24). On the IRF-PAI, the assessor is directed to list the most complicating comorbid conditions but they are also trained to first list the ones that will result in a payment adjustment. Hence, the reported comorbidities may under-represent conditions being actively treated or may not reflect those causing the greatest medical problems.

The MDS, on the other hand, gives a list of diseases and infections commonly found in long-term care cases (I.1 and I.2) and asks the assessor to identify if a disease or infection is related to the resident's current ADL status, cognitive or behavioral status, medical treatments, need for monitoring, or risk of death. This list is not as specific as ICD-9 codes, which also can reveal the location or severity of a condition as well as the type of illness. In addition, the MDS provides an open-ended item to list two other current diagnoses and ICD-9 codes (I.3.). The 3.0 MDS

(6) This was also a recommended IRF-PAI change so that this information would be collected for IRF patients.

version proposes to replace these items with ICD-9 codes, which will make these three forms more consistent in terms of conditions assessed. Severity is currently measured on the MDS in a general overall question (Item J.5.) rather than specific to an individual ICD-9 code.

e. Complicating Conditions: The tools vary in how they identify complicating conditions. Each of them list different conditions, or where they list the same type of complicating condition, it is often measured differently. For example, all three tools measure whether the patient is in pain, has skin conditions, such as ulcers or wounds, and their respiratory status. These are each important measures because they can reduce a patient's ability to participate in treatment independent of the severity of the primary illness. This changes the expected outcomes and levels of treatments that can be provided during a stay. As a result these have the potential to be important risk adjusters when examining case mix differences within and across settings. They also may explain differences in site of care used or level care needed.

i) Pain items: The IRF-PAI uses a ten-point scale that notes the worst level of pain documented on any shift during the prior three-day assessment period. In contrast, both the MDS and the OASIS document the frequency of the pain but use different measures. The MDS notes (M0420) whether a) no pain or not interfere with activity or movement, b) less than daily, c) daily, but not constant d) all the time or M0430) is intractable. OASIS captures this in two items. J2a asks frequency (no pain, less than daily, daily) and J2b asks intensity (mild, moderate, excruciating). While the "excruciating" OASIS item may be comparable to the MDS "intractable" item, this has not been tested. Further, the two "daily" responses may not be measuring the same level of frequency as one identifies the pain as "daily, but not constant" and the other tool is silent on whether the daily pain is constant or not.

ii) Skin conditions: Skin conditions are sometimes considered a quality indicator although it is difficult to determine when a skin condition actually began developing, and therefore, difficult to use as a quality measure.[6] Each tool measures the type and degree of skin conditions with different items.

The MDS section M addresses skin problems through a set of six items identifying:

1. Number of ulcers at each of the four stages during last seven days and as identified through a body exam,
2. Whether each ulcer identified in #1 is a pressure or stasis ulcer,
3. History of unresolved ulcers in last 90 days,
4. Other types of skin problems in last seven days including abrasions/bruises, rashes, desensitized skin spots, tearing, and surgical wounds
5. Whether any skin treatments were provided in the last seven days including special chairs, turning programs, nutritional or wound care programs, or other preventive activities
6. Series of foot and nail problems during the last seven days.

The OASIS asks a different series of questions about the current skin condition and whether the patient has a wound or ulcer (M0440s), the number and stages of ulcers (M0450), and the stages

and status of: a) pressure ulcers (M0460s), b) stasis ulcers (M0470s), and c) surgical wounds (M0480s). Measures are only reported on the most problematic ulcers rather than all ulcers as in the MDS.

The IRF-PAI is similar to the OASIS in that it asks for information on the worst ulcer rather than all ulcers but it also measures the severity of the ulcer based on width, exudation amount, and tissue type using the PUSH tool. Effectively, it is the most specific documentation of skin conditions although only for the worst conditions, unlike the MDS, which measures the severity of each ulcer.

This is a growing field of science that has been developing nurse certification and standards of practice for effective wound care management. One way to standardize the elements used to measure ulcers and skin conditions is to compare the current tools with the measures currently being developed in this field. Skin care treatment is an important preventive measure since patients with ulcers range from those with multiple, infected wounds needing treatment in an acute care setting, such as a long-term care hospital to wound management for a homebound person who may need debridement through an outpatient therapy setting, a fairly non-threatening condition that has the potential to become more severe.

iii) Respiratory conditions: Like the skin conditions, respiratory conditions are often complicating conditions in the more severely ill populations as well as primary diagnoses. Certain hospitals, like long-term care hospitals and some rehabilitation hospitals have specialized programs to treat patients with these conditions, including the use of special equipment such as ventilators. While ventilator weaning may occur in the LTCHs and some IRFs, nursing facilities may provide ventilator management for longer term residents. More information than whether a patient is on a ventilator is needed to distinguish between these levels of care.

Less intensive respiratory conditions are also noted on these forms. On the IRF-PAI, three questions currently exist that document whether the patient is short of breath in exertion, at rest, or has a weak cough or difficulty clearing airways at admission and discharge (N/Y items). The OASIS has two items (M0490 and M0500), which ask whether the patient is short of breath at different stages of exertion and whether they are using respiratory treatments such as oxygen, ventilators, or continuous positive airway pressure. While the OASIS M0490 is more specific than the IFPAI items 48-50, they could be crosswalked to answer whether the patient is short of breath under exertion or at rest or not at all. These questions target the healthier patient who has respiratory complications but not severe impairments.

The MDS records shortness of breath in items J1 (problem conditions) where it is noted if a patient has shortness of breath during the last seven days or an inability to lie flat due to shortness of breath. It also documents under section P whether the resident has had ventilator support or suctioning over the last 14 days.

f. <u>Special Equipment/Treatment Needs:</u> The third major domain under medical issues relates to the use of special equipment or special treatments. The IRF-PAI documents the smallest number since these patients are the healthiest. The IRF-PAI measures difficulties with swallowing (item 27) and whether the patient requires modified food consistency or supervision or the use of a tube or parenteral feeding either partially or completely for sustenance.

The OASIS documents tube/PTN feeding needs in two different sections. Under ADLs, (M0710) measures the level of ability to feed oneself (prior and current) and identifies whether assistance is needed and level of assistance (setup, supervision, modified diet). Responses range from able to feed self independently to unable to take in nutrients orally or by tube feeding. In addition, M0810 asks whether the patient needs assistance with equipment, including feeding, oxygen or IV/infusion therapy. It specifically refers to ability, not willingness and rates the patient on a four-point scale. A similar question is asked of the caregiver's ability to manage the three types of equipment (M0820). In addition, OASIS uses a three-point item to measure the patient's ability to take oral medications (M0780), inhalant/mist medications (M0790), and injectable medications (M0800). In complement to these medical issues, the OASIS also asks about the patient's prior and current abilities to complete their Instrumental Activities of Daily Living (IADLs).

The MDS documents similar issues such as the nutritional approach (K.5.) and the proportion of intake associated with tube use (K6.) but it also documents much more medically intensive needs than the IRF-PAI or OASIS. J1 lists a set of problem conditions that range from weight changes to dehydration to shortness of breath, edema, fever, internal bleeding, lung aspirations in the last 90 days to hallucinations and delusions. In addition, section P asks whether the patient has received any of a set of treatments in the last 14 days, including chemotherapy, dialysis, IV medications, ostomy care, radiation, suctioning, tracheostomy care, transfusions, ventilator or respirator support, Alzheimer's special unit services, hospice, respite care, or alcohol and drug treatments (P1). In addition, therapies in the last seven days are tracked, including physical, occupational, speech, respiratory, and psychological therapies (P2). The MDS also tracks whether the nursing staff has provided at least 15 minutes per day of range of motion or mobility training during the last seven days. Other items in the P section document whether, and how frequently, restraints have been used; frequency of physician visits and orders in the last 14 days, and whether the patient is likely to be discharged within a 90-day period. These items target services for both the medically complex patient and the long term, frail patient.

g. Functional Limitations: All three instruments contain items to assess functional status, although each tool relies on different items to elicit this information (Table 2.3). These three instruments are similar in their shared clinical concern with functional improvement. Despite these similarities, the information collected is not entirely comparable across the three instruments.[18]

The FIM instrument sums two subscales measuring the motor and cognitive independence levels. These scales are based on seven-level ratings of independence in each item including, eating, grooming, bathing, dressing, toileting, sphincter control, transfers, and locomotion. The FIM also measures comprehension, expression, and cognitive items such as social interaction, problem solving, and memory. All items are scored for their highest levels of dependence (from total independence to total dependence with varying levels of assistance between level 1 and 7) during the three prior days for an admission and discharge assessment.

The MDS uses a different set of functional items with a different scoring system. It measures functional impairment using a five level measure of ADL support provided during the last 7 days of the assessment period (no support, setup only, 1 person assist, 2+ assist, or did not occur) in

bed mobility, transfer, walking, locomotion, dressing, eating, toilet use, personal hygiene, and bathing.

The OASIS uses different scales to measure functional impairment at two points in time – prior to admission and currently (on the assessment day). It uses eight ADL measures with a four-level response scale. The four levels measure ability and level of assistance needed. They range from no assistance needed to materials must be placed within reach to individual assistance needed to total dependence on someone else.

While each of these tools are well-respected in their respective fields, little work has been conducted to apply them across settings. In recent years, several studies have tried to compare the functional measures on these three tools. They differ in whether they modified the tools, created a crosswalk using statistical methods, such as Rasch analysis, or collected data on one sample using multiple forms.

Table 2.3: Functional Measures Used on IRF-PAI, MDS, OASIS Tools

	IRF-PAI	MDS 2.0	OASIS
Assessment Period	Last 3 Days	Last 7 days	Current
Self-Care	39A. Eating	G1.h. Eating – how resident eats and drinks. Includes intake of nourishment by other means.	M0710. Feeding or eating – ability to feed self meals and snacks.
	39B. Grooming	G1.j. Personal hygiene – how resident maintains personal hygiene, including combing hair, brushing teeth, shaving, applying makeup, washing/drying face, hands, and perineum	M0640. Grooming – ability to tend to personal hygiene needs (i.e. washing face and hands, hair care, shaving or make up, teeth or denture care, fingernail care
	39C. Bathing	G2. Bathing – how resident takes full-body bath/shower, sponge bath, and transfers in/out of tub/shower	M0670. Bathing – ability to wash entire body excludes grooming (washing face and hands only)
	39D. Dressing-Upper	G1.g. Dressing – how resident puts on, fastens, and takes off all items of clothing, including donning/removing prosthesis	M0650. Ability to dress upper body – including undergarments, slacks, socks or nylons, shoes
	39E. Dressing-Lower	G1.g. Dressing – how resident puts on, fastens, and takes off all items of clothing, including donning/removing prosthesis	M0660. Ability to dress lower body – including undergarments, slacks, socks or nylons, shoes.
	39F. Toileting	G1.i. Toilet use – how resident uses the toilet room; transfer on/off toilet, cleanses, changes pad, manages ostomy or catheter, adjusts clothes	M0680. Toileting – ability to get to and from the toilet or bedside commode.

Table 2.3: Functional Measures Used on IRF-PAI, MDS, OASIS Tools (continued)

	IRF-PAI	MDS 2.0	OASIS
Assessment Period	Last 3 Days	Last 7 days	Current
Sphincter Control	39G. Bladder	H1.b. Bladder continence – control of urinary bladder function, with appliances, or continence programs, if employed	M0520. Urinary incontinence or urinary catheter presence M0530. Urinary incontinence frequency
	39H. Bowel	H1.a. Bowel continence – control of bowel movement, with appliance or bowel continence programs, if employed	M0540. Bowel incontinence frequency
Transfers	39I. Bed, Chair, Wheelchair transfers	G6. Modes of transfer – bedfast all or most of time, or bed rails used for bed mobility or transfer	M0690. Transferring: ability to move from bed to chair, on and off toilet or commode, into and out of tub or shower, and ability to turn and position self in bed if patient is bedfast
	39J. Toilet transfers	G1.i. Toilet use – how resident uses the toilet room; transfer on/off toilet, cleanses, changes pad, manages ostomy or catheter, adjusts clothes	M0690. Transferring: ability to move from bed to chair, on and off toilet or commode, into and out of tub or shower, and ability to turn and position self in bed if patient is bedfast
	39K. Tub, shower transfers	G2. Bathing – how resident takes full-body bath/shower, sponge bath, and transfers in/out of tub/shower	M0690. Transferring: ability to move from bed to chair, on and off toilet or commode, into and out of tub or shower, and ability to turn and position self in bed if patient is bedfast
Locomotion	39L. Walk/wheelchair locomotion	G1.e. Locomotion on unit – how resident moves between locations in his/her room and adjacent corridor on same floor. If in wheelchair, self-sufficiency once in chair	M0700. Ambulation/ Locomotion – ability to safely walk, once in a standing position, or use a wheelchair, once in a seated position, on a variety of surfaces
	39M. Stairs locomotion	G1.f. Locomotion off unit – how resident moves to and returns from off unit locations. If facility has only one floor, how resident move to and from distant areas on the floor. If in wheelchair, self-sufficiency once in chair	

Table 2.3: Functional Measures Used on IRF-PAI, MDS, OASIS Tools (continued)

	IRF-PAI	MDS 2.0	OASIS
Assessment Period	Last 3 Days	Last 7 days	Current
	39N. Comprehension – Auditory – Visual – Both	C6. Ability to understand others – understanding verbal information, however able	M0390 Vision impairment M0400. Hearing and ability to understand spoken language – in patient's own language with hearing aids if the patient usually uses them
	39O. Expression – Vocal – Non Vocal	C4. Making self understood – expressing information content – however able	M0410. Speech and oral (verbal) expression of language – in patient's own language
Social Cognition	39P. Social interaction		
	39Q. Problem solving	B4. Cognitive skills for daily decision making – made decision regarding tasks of daily life	
	39R. Memory	B2. Memory – recall of what was learned or known B3. Memory/recall ability – that resident was normally able to recall during last 7 days	M0560. Cognitive functioning – patient's current level of alertness, orientation, comprehension, concentration, and immediate memory for simple commands

CMS funded research during the late 1990s to develop a "unified" instrument that measured functional impairments across all three post-acute settings – IRFs, SNFs, and HH services. This tool, the MDSPAC, modified the MDS incorporating information from both the MDS and the MDS-home care. It differed from the FIM in how data were collected (input from multiple respondents, including all caregivers in addition to patient and family), the scoring system (multiple dependent episodes instead of just lowest dependency level during the three day look back period) and scoring range from 0 to 6 in reverse order of the seven-point scale used in the FIMs. In addition some of the scoring for ADL assist codes were changed to incorporate whether the patient was weight-bearing.[19] The study had mixed results. A factor analysis on items combined from both instruments found that the original MDS items did not load onto the same factors as the FIM's, while the revised MDS items had higher agreement with the FIMs items. However, the greatest agreement was achieved when the calibration teams actually collected data on the same patient using both forms. Despite this, when the team mapped the MDSPAC and the FIMs items into the IRF payment cells, the classification agreement levels were low. As a result, CMS opted to retain the FIM for use in the IRF.

Another study used the current MDS to develop a pseudo-FIM score for patients in IRFs and SNFs.[20] This study used an expert panel to crosswalk the FIM and MDS functional items. Data were collected in six nursing homes on rehabilitation patients. The FIM was administered by

therapists and nurses while the MDS was administered by nurses as is currently the practice. MDS levels were rescaled based on the TEP input to create pseudo-FIM measures. Ten FIM items had corresponding MDS items (eating, transferring bed to chair, toileting, bathing, grooming, bowel and bladder incontinence, problem solving, expression, and comprehension). Three FIM items were judged as not available on the MDS (climbing stairs, tub/shower transfer, toilet transfer). The TEP also disagreed on whether the walk/wheelchair FIMs item was equivalent to the MDS locomotion item since the latter incorporated distance. The cognitive FIMs items appeared to be related to several dichotomous MDS items so they and the three FIMs items without an MDS equivalent were dropped from the analysis. The results showed high levels of agreement (no significant difference in mean scores) on 8 items (eating, transferring, toileting, dressing, grooming, bladder control, memory, and problem solving. In testing the subscales, the researchers found a bias where subscales were higher (less impaired) for FIMs than pseudo-FIM on the motor and ADL subscales and lower on the cognitive subscales suggesting a ceiling and floor effect on certain items in the FIMs. This is consistent with earlier work which found there were no major ceiling effects, but certain patients those with lower level amputations and major multiple trauma groups, were unable to climb stairs and therefore their ratings were misleading. Similarly, patients tended to score higher at admission in eating, grooming, and the five cognitive items suggesting small ceiling effects.[21]

While the FIM serves as a valuable assessment instrument in its inpatient rehabilitation facility setting for which it is intended, it lacks sufficient variation in functional status in order to be used across the range of PAC settings.[22] The FIM primarily defines the distinctions within the lower levels of functioning where individuals with the greatest level of impairment score. The resulting concentration of scores at the higher end of the FIM scale constitutes a ceiling effect.[23]

While these tools are effective with the populations they were intended to address, they may mask variability when applied to other patients. As noted by Steinman et al, the floor and ceiling effects could be addressed by adjusting the definition of an item so that more people could be measured on it. Application to populations that are either too disabled or too able bodied could be misleading in that it would appear no functional gain occurred. Items would need to be added that would reflect much more difficult and much less difficult activities.[21]

An alternative approach has been used by Kramer et al., in which all three instruments have been mapped back to the Barthel (modified because stairs are not included on the MDS). In one study, IRF and SNF patients with hip fracture and stroke were assessed at PAC admission using the Barthel Scale, demonstrating substantially higher (more independent) Barthel scores in IRF patients than SNF patients.[24] More recently, the IRF-PAI, MDS, and OASIS scores for stroke patients were mapped to the Barthel Index, demonstrating the greatest independence among HH patients, moderate independence among IRF patients, and least independence among SNF patients.[25] This mapping to the Barthel Index, however, requires further validation.

More work needs to be done to obtain consensus and test the comparability of these measures.[26] Some work is underway collecting information using multiple tools on one set of patients. One study is collecting information on 600 patients using the FIMs, MDS ADLs, and the OASIS functional measures to create a crosswalk that can be used across settings.[27] Other researchers are interested in developing item banks that provide measures independent of setting and build

on the WHO ICF framework.[23] These types of studies can provide the tools needed to compare patients' functional limitations independent of site of care.

C. MISSING ELEMENTS FROM TOOL

The above discussion highlighted some of the issues in trying to standardize the types of information collected on PAC admissions using the current tools. The three existing tools are only used with the IRF, SNF, and HH admissions. However, as shown in Figure 1, these same types of cases may be discharged to long-term care hospitals (LTCHs). Often these cases are more medically complex than those admitted to IRFs but not always since the hour therapy rule has no intensity requirement. Both types of hospitals may wean a patient from a ventilator or provide rehabilitation services, depending on the region of the country. However, LTCHs do not currently submit any assessment data. As a result, they do not consistently document health or functional limitations in their patients.

A recent RTI study collected assessment tools from a survey of LTCHs to identify the types of factors monitored to determine treatment needs.[7] Many of the factors were similar to the MDS special treatment section. Items documented the use of ventilators, types of dialysis machines (peritoneal or hemodialysis), oxygen therapy, suctioning and tracheostomies. However, they also collect more objective measures of patient intensity, such as respiratory rates, stability of hemodynamic measures and other factors that measure intensity and the need for higher level monitoring or other services. This is particularly important when considering levels of care and determining whether a ventilator patient on certain drugs belongs in an ICU, LTCH, or IRF.

LTCHs varied in their collection of function items. Some LTCHs collected FIMs items while others used measures more similar to the MDS measures. Most were measuring performance rather than assistance provided. However, there appeared to be no standard measure of function across LTCHs.

These clinical and functional factors are important for many reasons. In addition to patient management, they are also useful to the Medicare program management. Many of these factors are used by the QIOs to assess appropriateness of acute hospital admissions, including LTCH admissions. Their current criteria are based on private sector efforts to develop "level of care" definitions. These systems use clinical measures such as comorbidities, blood and oxygen levels and equipment needs such as tubes, drips, and ventilator weaning or management. The ventilator items are particularly important because different levels of medical attention are needed by patients who are on a ventilator and being weaned versus the one who is being managed. The current indicator in the MDS does not distinguish between the two levels of intensity for patients on ventilators. These private sector efforts are similar to the RAI suites that have grown out of the MDS, but again, the measures have their basis in the medical rather than long-term care field.

D. CONCLUSIONS

The ultimate long-term goal of this effort is to create a minimum data set that collects standardized information on medical and functional status across settings, with additional components specific to each setting. Of utmost importance is the minimizing reporting burden. These data sets should be minimal so that the quality of care, outcomes, and case mix can be

tracked across settings without creating undue paperwork burdens. However, the limited research to date has shown that the current measures have each been designed to evaluate patients at a certain level of function. As stated in the beginning, IRF and HH admissions are likely to have greater strength and mobility than SNF admissions who likely have medical conditions and greater frailty, complicating any rehabilitation needs. And by definition, HH patients are likely to be healthier than SNF admissions in that the latter have to meet the needs for 24-hour nursing care.

Developmental work remains before we can create a unified assessment instrument. Logically, one would expect to develop new items that could build on each other to create better measures of a domain that can cross settings and provider levels of care. The World Health Organization has begun to build on this concept in their new taxonomy of disability and impairments. Through multiple governments and collaborating centers, the WHO is developing a 3,000 plus item pool to classify human functioning at the body, person, and social interaction levels.[28] When ready, this taxonomy can provide a set of measures that when used across the three levels can measure similar body impairments at different levels of physical and social limitations. It will be important, however, to have a minimum standardized set of items if one of the goals is to compare cases across an episode of care.

In the short term, the current system needs to be refined to allow continued use of the items included in the payment algorithms while still improving the ability to standardize case mix measures across settings for other purposes. The two most difficult areas to achieve consensus in are those measuring medical severity and functional impairments. By definition, the populations admitted to IRFs are "healthier" medically than those admitted to SNFs or LTCHs. They must be able to sustain three hours therapy per day, on average. Admissions at LTCHs are typically most medically complex as the average length stay is much longer than in other acute level settings and past work has documented the difference in severity between LTCHs and SNFs.

The work that is currently underway to build crosswalks in functional measures by using multiple forms on the same patients will be important in building an item bank that allows one to move across a wider continuum of measurement. These represent one type of effort to allow measurement across settings as beneficiaries continue to be discharged from one setting to the next lower level of care. In addition, research is needed to develop a refined item bank that can be used for uniform functional assessment.

E. REFERENCES

1 Gage, B. (Summer 1999). "Impact of the BBA on post-acute utilization." *Health Care Financing Review,* 20(4):103-126.

2 Gage, B., J. Kautter and E. Richter (October 2003). "Access To Outpatient Rehabilitation Therapy: Who Goes Where?" AARP Policy Paper.

3 Federal Register (August 1, 2003). Part III, p. 240, Table 5.

4 Gage, B., W.J. Bartosch, and D.S. Osber (August 9, 2005). "Medicare Post-Acute Care: Evaluation of BBA Payment Policies and Related Changes. Final Issues Report, CMS Contract No. 500-96-0006, T.O. #4.

5 Liu, K., C. Baseggio, D. Wissoker, S. Maxwell, J. Haley, and S. Long (2001). "Long Term Care Hospitals under Medicare: Facility-Level Characteristics." *Health Care Financing Review*, 23(2):1-18.

6 Gage, B., S. Bernard, R. Constantine, L. Greenwald, E. Root, C. Squire, J. Green, D. Osber, B. Gainor, and M. Steineman (July 2005a). "Development of Quality Indicators for Inpatient Rehabilitation Facilities." Final Report, Centers for Medicare and Medicaid Services, Contract No. 500-00-0024, T.O. #4

7 Gage, B., W. Bartosch, M. Leung, N. Pilkauskas, S. Hoover, J. Green (October 2005b). "Long Term Care Hospital PPS Monitoring and Evaluation" Phase II Report, Centers for Medicare and Medicaid Services, CMS Contract No. 500-00-0024, TO#20

8 Rice, P. (June 16, 2005). Testimony presented to Congress

9 Mor, V. (2005). "Improving the Quality of Long-Term Care with Better Information." *The Milbank Quarterly*, (83)3:333-364.

10 Katz, S. (1983). "Assessing Self-Maintenance: Activities of Daily Living, Mobility and Instrumental Activities of Daily Living." *American Geriatrics Society*, 31(12):721-726.

11 Prospective Payment Assessment Commission (1992). "Prospective Payment System for Medicare's Skilled Nursing Facility Payment Reform" Congressional Report, C-92-01, Washington, DC: ProPAC.

12 InterRAI (2005). InterRAI Integrated Suite of Instruments. http://www.interrai.org/section/view/?fnode=38. Accessed 1/22/06

13 Mahoney, F. I. and D. W. Barthel (1965). "Functional Evaluation: The Barthel Index." *Maryland State Medical Journal*, 14:61-65.

14 Carter, G., D. Relles, G. Ridgeway, and C. Rimes (Spring 2003). "Measuring Function for Medicare Inpatient Rehabilitation Payment." *Health Care Financing Review*, 24(5).

15 Haley, S. M., W. J. Coster, P. L. Andres, L. H. Ludlow, P. Ni., T. L. Y. Bond, S. J. Sinclair, and A. M. Jette (January 2004). "Activity Outcome Measurement for Postacute Care." *Medical Care*, 42(1 suppl):I-49–I-61.

16 IRFPAI Training Manual: The Inpatient Rehabilitation Facility-Patient Assessment Instrument Training Manual, U.B. Foundation Activities, 2001-2004, Buffalo, New York, April 2004.

17 Medicare Payment Advisory Commission (June 2005). "Report to Congress: Issues in a Modernized Medicare Program." Washington, DC: MedPAC.

18 Iezzoni L.I. and M.S. Greenberg (Spring 2003). "Capturing and Classifying Functional Status Information in Administrative Databases." *Health Care Financing Review,* 24 (3):61-76.

19 Buchanan, J.L., P.L. Andrews, et al. (Spring 2003). "An Assessment Tool Translation Study." *Health Care Financing Review*, 24(3),:45-60.

20 Williams B.C., B.E. Fries, and R.L. Warren (January 1997). "Predicting Patient Scores Between the Functional Independence Measure and the Minimum Data Set: Development and Performance of a FIM-MDS 'Crosswalk'." *Archives of Physical Medicine and Rehabilitation*, 78:48-54.

21 Stineman, M. G., J. A. Shea, A. Jette, C. J. Tassoni, K. J. ottenbacher, R. Fiedler, and C. V. Granger (November 1996). "The Functional Independence Measure: Tests of Scaling Assumptions, Structure, and Reliability Across 20 Diverse Impairment Categories." *Archives of Physical Medicine and Rehabilitation*, 77:1101-1108.

22 Jette A.M., S.M., Haley and N. Pengsheng (Spring 2003). "Comparison of Functional Status Tolls Used in Post-Acute Care." *Healthcare Financing Review*, 24 (3):13-24.

23 Jette A.M. and S. Haley (2005). "Contemporary Measurement Techniques for Rehabilitation Outcomes Assessment." *Journal of Rehabilitative Medicine*, 37:1-8.

24 Kramer AM, Steiner J, Schlenker RE, Eilertsen TB, Hrincevich CA, Tropea DA, Ahmad L, Eckhoff DG. "Outcomes and Costs after Hip Fractures and Stroke: A Comparison of Rehabilitation Settings." JAMA 277(5):396-404, 1997.

25 Kramer, AM, G. Goodrich, D. Holthaus, & A. Epstein. (2006). Study of Stroke Post-Acute Care and Outcomes: Draft Final Report. Chapter 4: Paterns of Post-Acute Care Following Implementation of PPS. Division of Health Care Policy and Research, UCDHSC, Aurora, CO (in preparation).

26 Clauser, S.B. and A.S. Bierman. (Spring 2003). "Significance of Functional Status Data for Payment and Quality." *Health Care Financing Review*, 24(3):1-12.

27 Mallinson, T. (2004). "Recovery of Function Across Rehabilitation Settings: An Equating and Effectiveness Study." Center for Rehabilitation Outcomes Research, Rehabilitation Institute of Chicago, NIDRR Grant H133A030807.

28 Ustun, T. B., S. Chatterji, N. Kostansjek, and J. Bickenbach (Spring 2003). "Who's ICF and Functional Status Information in Health Records." *Health Care Financing Review*, 24(3):77-88.

Chapter 3
Review of Uniform Assessment Items and Tools for Ambulatory and Hospital Settings

A. OVERVIEW

Providers, researchers, and health systems have explored a plethora of approaches to standardizing and improving the assessment of older adults. In this chapter, we will identify, from among these approaches, those domains that are commonly included in efforts to measure important patient characteristics, outcomes, and basic transitional care needs. We will also provide examples of some items that are used to measure these domains. We will then discuss some of the challenges that have been encountered in measuring these domains. Finally, we will review the content of some available assessment tools and approaches that have been proposed or tested in ambulatory and hospital settings. Although the list of measures and assessment tools is too extensive to review in its entirety, this review involved an attempt to identify the most widely used assessment instruments and measures outside of those tools used under Federal mandates, including instruments used for managed care and private initiatives.

This review will focus on those assessment domains and related items that might further the potential goals for a hospital discharge assessment instrument. As outlined in Chapter 1, a standardized assessment might be used to measure those patient characteristics that significantly influence outcomes as patients receive care in different settings over time. By assessing these characteristics, providers are better able to anticipate their patients' outcomes and improve their quality of care. Another goal of a standardized hospital discharge assessment might be to improve discharge care planning. By assessing those factors that influence the patient's discharge outcomes, care and support needs, providers might better tailor discharge recommendations to meet their patients' needs. A third potential goal for a standardized discharge assessment is to improve care transitions by identifying patient characteristics and care elements that allow subsequent care providers to best manage care. The remainder of this chapter will focus on domains, items within domains, and examples of existing assessment tools that have been identified, either in the medical literature or in expert opinion, as important for meeting one or more of these potential goals of a hospital discharge assessment instrument.

B. POTENTIAL DOMAINS

Multiple domains have been included across standardized patient assessments. The large number of recognized factors that contribute to important health outcomes[1, 2] can be divided into four groups—demographics and socio-economic status (SES);[3-10] function;[10-16] health behaviors and preferences;[7, 17-19] and diagnoses, conditions, and associated treatments.[15, 17, 18, 20-24]

1. Group I: Demographic and SES Domains

An individual's age influences his or her response to disease and chances of functional recovery, decline and death,[25-29] even after accounting for other patient characteristics.[8, 13, 17, 19, 24, 30-36] As a result, age is included in many assessment models or scoring systems.[11, 18, 37-39] Care guidelines and care recommendations for preventive and screening services often are targeted to particular age groups. In addition, age has been associated with differences in care quality and access.[40-42]

Gender and marital status have been associated with differences in health outcomes, may play some role in patient discharge choices, and are relevant for future medical care. Male gender may be protective for some outcomes,[43] but is associated with increased risk for others.[40] In addition to variations in disease incidence and prevalence, both quality of care and proclivity to report symptoms may vary by gender.[42, 44-47] Marital status may influence health states,[48, 49] choice of discharge setting,[50] and may provide relevant information for future providers. However, the relationship can be complex and interacts with gender and the health of the caregiver.[48, 50, 51] Measures of caregiver availability and social support may be more important for predicting future needs, access, and discharge planning.[37, 52, 53]

The importance of language for effective communication and health literacy is clear and is relevant for interpreting survey responses, scoring cognitive tests, and discharge planning.[54-56] The role of race and ethnicity, independent of language, is more complex. These factors continue to be associated with differential access, poor care transitions, care delivery and quality,[50, 54, 57-60] although care reports from some settings offer hope that these quality differences can be eliminated.[60-65] As such, race and ethnicity are relevant for monitoring disparities in outcomes. Race and ethnicity have also been associated with differences in health beliefs, behaviors and content of self-report.[56, 64, 66, 67] These associations point to possible utility for care planning, care transitions, and identifying best items for a standardized assessment. Finally, race may independently influence disease prevalence and outcomes, with various analyses showing differential disease progression and medication response, even after controlling for quality and other risk factors.[68-70] Pharmacology sometimes uses race as a crude proxy for applied genomic pharmacology.[71-73]

Education, health literacy, and income are important influences on health outcomes and access to services.[7, 17, 32, 74, 75] In addition, education and literacy influence patients' ability to comprehend and adhere to discharge instructions and to answer particular assessment items.[55, 76]

2. Domain Group II: Functional Status

Many widely accepted health models emphasize function over disease and define important disability as the functioning of individuals within their environment.[77] Function potentially clarifies the severity and impact of self-reported medical conditions in the geriatrics population.[78] Several guidelines advocate documentation of functional status on admission to hospital or physician practice.[79] Functional status is a broad dimension of health, and several scales are available for measuring functional limitation and disability. Functional measures can be divided into three large groups: measures of physical function, measures of the instrumental activities of daily living (IADL), and measures of the basic activities of daily living (ADL).[(1)] Although many tools have items from all of these domains, more recent scholarship treats them as separate functional domains, at least for purposes of measurement. We will discuss each group separately below.

[(1)] This grouping is used to facilitate practical review of existing instruments. The WHO International Classification of Impairments Disabilities and Handicaps employs a conceptual framework in which disease can lead to three general functional outcomes: impairment, disability, handicap. While we are aware of this model, it is not employed in this review.

Physical function or performance measures aim to capture underlying impairments and range from ability to accomplish gross physical activities such as climbing stairs[80] to measures of dimensions of physical activities such as strength, range of motion, and speed.[25, 30, 81, 82] The most commonly used measures focus on lower extremity strength, mobility and balance.[81] Physical function clearly influences health outcomes and gives the provider important information about a patient's current health status.[11, 30, 77, 83-89] Routine, standardized assessments of physical function can aid providers in recognizing declining health and items such as ability to climb stairs or walk are relevant for discharge planning and needs assessment to minimize resulting disability and handicap. Some measures are based on patient self-report[80, 90,91] and others are based on provider tests and observations.[92-96] The choice of self-report vs. provider-tested function should be considered in the context of the purpose of the measure: Self report captures the patient's view of his/her abilities, which is critical for care transitions and to examine outcomes over episodes of care; whereas providers may wish to record observed physical functions for care planning and monitoring. Table 3.1 gives two examples of item sets that emphasize physical function.

Instrumental activities of daily living (IADLs) and basic activities of daily living (ADLs)[97-100] are a mainstay of geriatric assessment and serve to describe an individual's ability to accomplish tasks needed for self-maintenance and independence in the community.[101] The ADL scales typically include items such as bathing, dressing, toileting, eating, transferring and continence. Several scales include ability to walk across a room with or without an assistive device. The Katz ADL, Barthel Index, and Functional Independence Measures were all created as observer reported ADL scales. In common practice, the Katz ADL items and Barthel Index are used as self or proxy reported items. IADLs/ADLs can be queried with different response scales "have difficulty," "need human assistance," or "receive help/human assistance." IADLs and ADLs reflect both underlying physical and cognitive limitations.[102-104] ADLs have been noted to have a hierarchical relationship for loss.[103, 105, 106] This hierarchical structure can be useful as providers review patient status. IADLs do not consistently demonstrate this property.[99, 105]

Because of their focus on an individual's ability to function independently in the community, IADLs/ADLs are important quality of care outcomes. In addition, IADL/ADL difficulty and requiring or receiving help are recognized and robust predictors of mortality, hospitalization, and nursing home use.[11, 17, 24, 105, 107-111] IADL disability also predicts outpatient and home care resource use. A common concern with ADL items is whether they are sufficiently sensitive to clinical improvement and change. However, the longer Functional Independence Measure (FIM), with its 7-point response scale, has shown more sensitivity to change.

Table 3.1: Illustrative Item Sets that Emphasize Physical Function

Name	Administration	Item Content
SF 36 Physical Function Scale	Self Report	*"The following items are about activities you might do during a typical day. Does your health now limit you in these activities?"* 1. ***Vigorous Activities,*** *such as running, lifting heavy objects, participating in strenuous sports* 2. ***Moderate Activities,*** *such as moving a table, pushing a vacuum cleaner, bowling, or playing golf* 3. ***Lifting or carrying*** *groceries* 4. ***Climbing several*** *flights of stairs* 5. ***Climbing one*** *flight of stairs* 6. ***Bending***, ***kneeling***, *or* ***stooping*** 7. ***Walking*** *more than a* ***mile*** 8. ***Walking*** *several* ***block****s* 9. ***Walking one block*** 10. ***Bathing or dressing*** *yourself."* [a] Response options for each item = *Yes, limited a lot; Yes, limited a little; No, not limited at all.*
Summary Performance Score, Lower Extremity[87, 96]	Observation	Scaled score combining: *Walking Speed* "the participant was asked to walk over a 4-meter course…Participants were instructed to stand with both feet at the starting line and to start walking after a specific verbal command. Timing began when the command was given. … subject could use a cane, a walker, or other walking aid, but not the aid of another person. The times to complete the first meter and the entire path were recorded. The test was repeated three times, twice at the woman's usual pace and once at her fastest possible pace. The speed of the faster of the two usual-pace walks is presented. The length of the walk expressed in meters divided by the time in seconds was used to calculate average walking speed." *Chair Stand* "participants were asked to sit with their arms folded across their chests in a straight-backed chair placed with its back against a wall, and then to stand up from the chair one time. If they were successful, they were asked to stand up and sit down as quickly as possible five times in a row. Timing commenced from the initial sitting position and ended at the final standing position at the end of the fifth stand." *Standing Balance* "Static balance is evaluated in three different, progressively more difficult stances…: (1) side-by-side: feet side by side, touching; (2) semi-tandem: side of the heel of one foot touching the big toe of the other; (3) tandem: heel of one foot directly in front of and touching the toes of the other foot. Each stance is progressively more difficult to hold. Women unable to hold a position for 10 seconds were not asked to attempt further stands."

[a] Note, this later item is typically considered an ADL, however, it is included here as part of the SF-36 physical function scale.

Table 3.2 summarizes the activities included in two common ADL assessment tools. Many consider the Katz scale as the most fundamental set. The Barthel Index uses a different response metric for various items and adds two mobility items, increasing the assessment range of the index.

Table 3.2: Comparison of Katz Scale and Barthel Index Activities

Instrument	Response scale	Activities Included
Katz	3 points: Independent Limited Assistance Extensive Assistance/ Cannot do	Eating/feeding Bathing Dressing Toileting Transfer, chair Continence
Barthel Index	Varies by item 0-5 for bathing grooming 0-10 for feeding dressing continence toilet use 0-15 for transfers mobility (from immobile to >50 yards)	Feeding Bathing Grooming Dressing Bowel continence Bladder continence Toilet use Transfer, bed to chair Mobility on level surface Stairs

The Lawton list of IADLs is shown below.

- Unable to use telephone
- Unable to take care of all shopping needs
- Unable to plan, prepare, and serve meals
- Unable to do any laundry
- Unable to maintain house alone (excludes occasional assistance)
  - unable to perform light daily tasks
  - cannot keep clean
  - needs help with maintenance
- Drive car or travel alone on bus or taxi
  - arranges travel by taxi
  - can use public transportation if assisted
- Not responsible for taking own medications
- Not manage financial matters independently
  - not manage day-to-day purchases

Despite the importance of functional status measures for patient well-being and function, most providers fail to assess or document them routinely during the course of care for older adults.

3. Domain Group III: Health Behaviors and Preferences

Tobacco use strongly influences health and the ability to recover from injury or illness.[17, 34] Screening for tobacco use and providing guidance and assistance with cessation are considered quality care processes.[79, 112] Knowledge of a patient's level of alcohol use is important and improves quality[79] because of the potential interaction of alcohol with many medications.[113] The

number of drinks a week can be directly queried. In addition, excessive alcohol use can be an important influence on overall health, safety, and adherence to medications.[114] Screening tools such as the Cut down, Annoyed, Guilty, Eye-opener (CAGE), Alcohol-Use Identification Test (AUDIT), and the Alcohol-Related Problems Survey (ARPS)[115] have been used to identify problem drinking in older adults.

The role of physical activity and diet in patient prognosis and improvement are clearly recognized,[116] and screening is considered important for high quality care.[79] In addition to IADL items addressing ability to prepare meals and mobility items discussed under functional status above, food access may be relevant for discharge planning. Likewise, dietary needs and allergies related to conditions below would be of importance.

Patient autonomy and control over health decisions is a recognized care value. Advance care plans and goals of care seek to formally document patient preferences for future interventions in the event the patient becomes unable to express preferences. Despite the seeming wisdom of such planning, providers infrequently devote resources to assisting patients with this process. Simple queries, whether part of a study intervention or mandated data collection, do not appear to improve the quality or frequency of completion of advance care directives.[117] It is important, however, to standardize transmission across care sites of those preferences that have been expressed and documented. The POLST, a one-page form, includes current treatment restrictions, treatment restrictions for future states, and decision-making. It has had overall success in one region for conveying advance care information about nursing home residents across providers.[118-122]

4. Domain Group IV: Diagnoses, Conditions and Associated Treatments

Active diagnoses requiring care management and follow-up are clearly important for discharge planning and care transitions. Medical diagnoses and conditions are recognized factors influencing death, hospitalization, functional improvement, and functional decline. The common medical diagnoses and conditions that are important for improving outcomes and discharge planning are dementia,[12, 15, 123-126] diabetes,[8, 127, 128] stroke,[19, 23, 127-129] mood disorder and other psychiatric diagnosis,[20, 130, 131] myocardial infarction or angina,[19, 132, 133] valvular heart disease,[17, 134] heart failure,[17, 135] other heart conditions,[23] limited vision,[23, 128] limited hearing,[23, 136] limited vision or hearing,[26, 133] hip fracture,[137] cancer,[22, 138] falls,[139, 140] and arthritis.[21, 127] While these are the most commonly named significant risk factors for decline or future utilization, multiple other diagnoses and conditions may be important for a given patient. These diagnoses and conditions would vary by individual and could include, but not be limited to, the list provided above.

Medications used to treat disease are a marker for active diagnoses, disease severity and health status.[141] Medication errors continue to be a common source of medical error. Such errors can stem from drug-to-drug interactions, from drug-disease interactions or from potential over- and under-utilization of certain classes of drugs.[142] Clear and accurate transmission of reconciled medication lists (admission and discharge lists with name, indication, dose; allergies) is central to safe care transitions. The Medicare Modernization Act offers an opportunity to systematically collect medication information and could be a powerful supplement to assessment data and quality improvement activities.[143, 144]

Administrative data are frequently used to identify diagnoses and conditions. These data typically have better specificity than sensitivity for capturing medical diagnoses and conditions.[145-148] A few exceptions are notable. Outpatient diagnoses may be assigned to justify diagnostic testing for some conditions leading to a decrease in specificity of administrative data used for filing claims.[149, 150] In addition, long-standing problem lists may lack specificity for detecting active disease and should be distinguished from other administrative data.

a. Under-detected Conditions and Syndromes With Potential Importance for Standardized Assessment Tools: Although administrative data have acceptable sensitivity for noting many conditions, under-detection is a significant problem for several important conditions and geriatric syndromes[151, 152] such as dementia,[153, 154] depression,[155-157] pain,[158-160] delirium,[161] and falls.[162] Medical record data and administrative data frequently demonstrate poor sensitivity for identifying these conditions because providers fail to systematically assess for their presence. This has important quality of care implications because these under-detected conditions are critical to the quality management of vulnerable older adults,[79, 163] may be caused by drug side effects or interactions,[164] are important for discharge planning and improving care transitions,[165] and are the conditions most likely to evidence quality deficiencies in care processes.[166] In addition, with the exception of falls, these conditions and syndromes have each been associated with variations in how patients report on preferences, other symptoms, and conditions. As such, measurement is particularly important if a discharge assessment tool will contain patient self-report items. The following paragraphs discuss some approaches to screening for these conditions.

b. Depression: Several depression screening tools have been tested and found to be valid. Most instruments screen for symptoms of possible depression. Because some symptoms such as disturbed sleep may reflect diseases other than depression, final diagnosis requires additional clinical identification that the symptoms relate to depression and that DSM-IV depression criteria are met. Some of the more commonly used self-report depression screeners are the Geriatric Depression Scale (GDS);[167, 168] the Center for Epidemiological Studies Depression Scale (CES-D);[169] the Primary Care Evaluation of Mental Disorders (Prime MD);[170, 171] the PRIME-MD 2-item;[172] and the Patient Health Questionnaire (PHQ-9).[131, 173-175] The Cornell Scale for Depression in Dementia (CSDD) and the Hamilton Depression Rating Scale (Ham-D) are among the more commonly used interviewer or observer rated tools for use in severely cognitively impaired or non-communicative patients. Both the CSDD and the Ham-D require considerable evaluator training to achieve acceptable reliability.[176, 177]

The CES-D and GDS are the more long-standing tools and represent a significant advance from older surveys that asked for somatic symptoms such as palpitations. The CES-D consists of 20 items, has acceptable psychometric properties in adults, but has had limited acceptability to providers. The GDS was developed specifically for older adults and also has acceptable psychometric properties, but has not been widely adopted outside of the Geriatrics community. Thirty-, fifteen-, twelve-, and five-item versions of the GDS are available. The 15-item GDS shows almost equivalent sensitivity and specificity to the longer 30 item version. One study that compared the CES-D to the GDS in 130 community-dwelling adults age 60 or older found comparable performance. Sensitivity and specificity were 0.92 and 0.81 for the GDS-15 and 0.92 and 0.87 for the CES-D.[178] Specificity is typically significantly lower in sicker populations. In a nursing home sample that was limited to persons with mild or no cognitive impairment,

comparable sensitivity and specificity were 0.88 and 0.62 for GDS-15 and 0.74 and 0.70 for CES-D.[179]

The PRIME-MD, PRIME-MD 2-item, and PHQ-9 are a related group of depression screening tools. The PRIME-MD is the longest. It is a 27-item questionnaire with follow-up clinical interview. The PRIME-MD is considered a gold standard diagnostic tool, but its length prevents widespread implementation as a screening tool. The shortest is the PRIME-MD 2-item. Its two items have a high sensitivity (0.96) but a low specificity (0.57) for detecting depression. The developers acknowledge that this level of specificity means that at least 35 out of 100 persons screened would have a false positive result. This would seem to be too high for a publicly mandated assessment tool. The PHQ-9 is a self-report version of the PRIME-MD with a reported sensitivity of 0.88 and specificity of 0.88. Its items directly relate to DSM-IV criteria for diagnosing depression. The PHQ-9 is increasingly being used across health care settings,[180, 181] and has been shown to be sensitive to change.[173] It has been used effectively in practice-based interventions to improve primary care of depression in older adults.[175, 182-184]

In addition to the PRIME-MD 2-item, other one- and two-item screeners have been explored. One study found that the Yale Depression single-item screen "Do you often feel sad or depressed" had comparable sensitivity and specificity to the GDS (0.69 and 0.90 for Yale Depression; 0.54 and 0.93 for the GDS)[185] but some have suggested that these single-item approaches may not perform as well in older populations.[186] Moreover, these abbreviated items would be unlikely to show sensitivity to change and therefore might fail to meet the important need for serial assessments to monitor response to therapy.[187] Table 3.3 describes selected self-report depression screening instruments.

Table 3.3: Examples of Self-Report Depression Screening Instruments

Instrument	Response metric	Questions
PRIME-MD	Yes No	During the past month have you been bothered a lot by 1. Little interest or pleasure in doing things 2. Feeling down, depressed, or hopeless
PHQ-9	Not at all Several days More than half the days Nearly every day	Over the last 2 weeks, how often have you been bothered by any of the following problems? 1. Little interest or pleasure in doing things 2. Feeling down, depressed, or hopeless 3. Trouble falling or staying asleep, or sleeping too much 4. Feeling tired or having little energy 5. Poor appetite or overeating 6. Feeling bad about yourself — or that you are a failure or have let yourself or your family down 7. Trouble concentrating on things, such as reading the newspaper or watching television 8. Moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual 9. Thoughts that you would be better off dead or of hurting yourself in some way

Table 3.3: Examples of Self-Report Depression Screening Instruments (continued)

Instrument	Response metric	Questions
GDS-15	Yes No	1. Are you basically satisfied with your life? 2. Have you dropped many of your activities and interests? 3. Do you feel that your life is empty? 4. Do you often get bored? 5. Are you in good spirits most of the time? 6. Are you afraid that something bad is going to happen to you? 7. Do you feel happy most of the time? 8. Do you often feel helpless? 9. Do you prefer to stay at home, rather than going out and doing new things? 10. Do you feel that you have more problems with memory than most? 11. Do you think it is wonderful to be alive now? 12. Do you fell pretty worthless the way you are now? 13. Do you feel full of energy? 14. Do you feel that your situation is hopeless? 15. Do you think that most people are better off than you?

c. Pain: Pain is a prevalent condition in older adults. For patients with pain, appropriate assessment and tracking of pain is an important element of care.[79, 188] It is also an important quality of life outcome and relates to other outcomes.[140] Because of the need to track pain symptoms, treatment response, and the need to determine if pain is new or worsening, a standardized measure should be considered for inclusion in information conveyed at care transitions. As an indicator of the importance of serial assessment, standardized pain measurement is sometimes referred to as the "fifth vital sign."[189-191] Several different generic (i.e., not condition-specific) pain scales have been developed and used across care settings.

The most commonly used pain intensity scales are the 0 to10 Numeric Rating Scale (11-point NRS) and Verbal Descriptor Scales (VDSs). The 11-point NRS is commonly used in hospital settings. It can be administered verbally, asking the patient to rate pain with 0 indicating "no pain" and 10 indicating "worst pain can imagine." The preferred mode of administration in older adults or populations with mixed communication abilities is face-to-face with both verbal questioning and visual display of the NRS scale in large print. An example of a NRS is shown in Figure 3.1. Other NRS number lines might show only the bottom and top anchor. Verbal descriptor scales offer a variety of response metrics. A common VDS is the one included in the McGill Short Form questionnaire and is listed in Figure 3.1. A different VDS commonly used independently in the literature is also shown. Other tools, such as the faces rating scale, may have poorer performance in population sub-groups such as cognitively impaired.[188, 192]

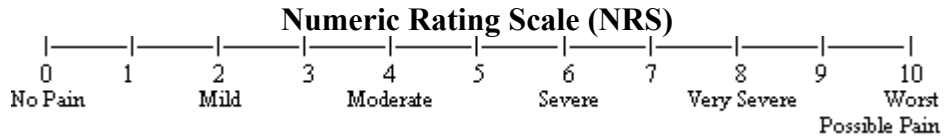
Longer pain multi-dimensional assessment tools also consider frequency of pain, location of pain, and effect on function. The original McGill Pain Questionnaire is a standard in the field, using body maps to identify pain location, word groupings to describe pain, and items for rating pain. It also includes items on history, medications, etc. The authors report that it takes 15-20 minutes to initially complete and the time decreases to 10 minutes or less with interviewer experience.[193] Shorter forms of the McGill have been proposed but these remain time consuming. The Geriatric Pain Measure is a shorter multidimensional pain assessment designed

to evaluate older adults. It has been shown to have convergent validity with the McGill (pearson's r = .63, p < .0000). It includes 24 items that address pain intensity, disengagement, pain with walking, and pain with activities.[194]

Figure 3.1: Examples of Numeric Rating Scales and Verbal Descriptor Scales

Numeric Rating Scale (NRS):

"On a scale of zero to ten, where zero means no pain and ten equals the worst possible pain, what is your current pain level?"



Verbal Descriptor Scales (VDS)-- two examples:

McGill VDS

Rate the overall intensity of your pain

0. *No pain*
1. *Mild pain*
2. *Discomforting*
3. *Distressing*
4. *Horrible*
5. *Excruciating*

Other commonly used VDS

Have you had pain over the past two weeks? (or Do you have pain now?)

0. *No*
1. *Yes*

If Yes, ask patient to describe the intensity of the pain at its worst

1. *Mild*
2. *Moderate*
3. *Severe*
4. *Very severe, horrible*

d. Cognitive Impairment: Cognitive impairment influences discharge location[195] and rehabilitation outcomes[196] and is important for discharge planning and safety issues during care transitions. Some quality indicator sets advocate documentation of cognitive status on admission to the hospital or a new physician practice.[79] The US Preventive Task Force, however, has failed to find sufficient evidence to support routine screening.[112] Multiple tools have been developed to aid in cognitive screening and assessment. The most commonly used and accepted cognitive assessment is the Mini Mental State Exam (MMSE).[197] Its 20 items assess orientation, memory (recall), calculation/task memory, language, and visuospatial skill. It must be administered face-to face and requires approximately 10 minutes to complete. For those providers who do screen for cognitive impairment, the MMSE has gained wide clinical acceptance. It correlates well with more in-depth gold standard neurocognitive testing by trained clinicians, although some cognitive domains are not assessed by the MMSE. It may be most lacking in its ability to identify early cognitive impairment and early loss of executive function, even after using accepted norms to adjust for education and age. Nonetheless, the MMSE has become the standard assessment against which shorter assessments are typically compared.

Clock drawing has been proposed as a simple screen for early cognitive impairment or loss of executive function that may be missed by the MMSE.[198] It must also be administered face-to-

face. The Clock draw test may have some value as an independent screen for care planning because it may identify persons at risk for future institutionalization.[199] However, it may not have equal validity in different cultural groups[200] and even when coupled with MMSE is not sufficient for screening for all cognitively-related safety concerns.[201]

The Mini-Cog combines 3-item recall with the clock draw.[202] A score of 0-2 suggests dementia. When "naïve" interviewers use the Mini-Cog, their performance has been found to be comparable to trained interviewers.[203] Sensitivity and specificity for classifying as "some" versus "no" impairment was similar to the MMSE in one community based sample, sensitivity = 76% vs. 79%, and specificity = 89% vs. 88% for dementia.[204]

Researchers have developed abbreviated tools that can be administered by telephone. One such tool uses 3-item recall (apple, table, penny) and 3-item temporal orientation (day week, month, year) items to screen for cognitive impairment. Greater than 3 errors has sensitivity (89%) and specificity (88%) for classifying as dementia.[205] Other cognitive screening tools that can be administered by telephone include the 7-item Blessed Orientation Memory Concentration Test (BOMC)[206] and the 10-item Short Portable Mental Status Questionnaire (SPMSQ).[207] For these shorter instruments, brevity and ease of administration come at the expense of floor and ceiling effects and ability to identify executive dysfunction. The Callahan and BOMC tools offer limited ability to measure change over time. The SPMSQ does allow some differentiation of levels of impairment. The Telephone Interview for Cognitive Status (TICS), has been more recently created for telephone screening for research purposes. It has a proxy version and does allow differentiation of levels of impairment. It requires about the same amount of time to complete as the MMSE, and may have fewer ceiling effects.[208] However one recent, small study suggests low sensitivity (0.38) for detecting age associated memory impairment.[209]

e. Delirium: Delirium is another under-detected health condition that is associated with poor health outcomes across settings,[210-213] is a marker of hospital quality,[214] and is important to detect in order to improve care transitions. It is associated with significant increase in current and future healthcare utilization.[215, 216] Detection is complicated by the fact that delirium may be superimposed on dementia or may occur in persons with no prior history of cognitive impairment. The Confusion Assessment Method (CAM)[217] is an accepted gold standard screening tool for delirium. It is completed after conducting a structured cognitive exam such as the MMSE, and requires significant staff training to achieve acceptable sensitivity or reliability.[218]

f. Falls: A history of falling is related to important patient outcomes including future falls, hip fracture, recovery, mortality, and self-imposed limitation in mobility because of a fear of falling.[128, 140, 219-223] Discharge planning might be improved by identifying those patients with fall history.[139] This can be accomplished by a simple query regarding number of falls in past year, classifying responses as 0,1, >1. In addition, it is important to note that some persons who limit their activity because of fear of falling do not have a prior history of falls.[224] However, measuring fear of falling may require a more complex set of items.[225]

C. CHALLENGES FOR MEASURING DOMAINS

The preceding discussion presents domains with strong support for inclusion in standardized assessments. There is less consensus regarding their priority for inclusion or the best specific items for measuring these domains across population groups and across care settings. Unfortunately, the evidence base for identifying "best items" varies widely in sampling, the number and combination of considered domains and items, and the approach to querying those important patient domains and outomes.[226, 227] This variation yields many apparently contradictory findings.[228] Unfortunately, very few studies have been conducted in older adults to directly compare the performance of different assessment instruments or the best combination for covering the full range within domains.

As indicated by the date range of the references above, most of these domains have been recognized and confirmed as important predictors for many years. The current challenge is selecting from among the potentially relevant domains those that are most appropriate for the purposes of this assessment, and then identifying the best ways to measure each domain to meet these needs. Once a domain is targeted, a further challenge arises from the need to include items and scales that are relevant for the entire Medicare PAC population. Covering the full range within a domain to avoid floor and ceiling effects can be extremely burdensome. Approaches such as those presented in Chapter 5 offer promise for minimizing burden while assuring precise measurement across the population. Developmental work might require the collection of data to allow better comparisons of performance in diverse populations of older adults in order to address the potential challenges to item validity and relevance outlined below.

The potential challenges to validity and feasibility of the assessment include achieving agreement across assessors in actual care settings, identifying how items will be collected, minimizing burden, achieving acceptability to providers and patients, accounting for item performance across population subgroups, accounting for the effect of variations in item wording on the relationship of items to each other and to the underlying construct, and translation of data into improved care. Each of these is discussed briefly below.

1. Agreement Across Assessors (inter-rater reliability)

Reliable information and reporting are fundamental to achieving assessment validity. Reliability and validity must be considered in the context of both efficacy (performance in ideal circumstances), and effectiveness (performance in actual conditions).[229] Ideal or research collection of items may differ from collection in the real world because research teams are typically well trained for and charged with focusing on a limited set of measurement tasks. On the other hand, modern practice time pressures may limit provider attention to assessment, particularly if a significant amount of training or clinical evaluation time is needed. This makes it difficult to achieve reliability when collecting a large number of items in active clinical settings. In addition, users may lack motivation to invest in accurately assessing mandated items, particularly if they do not perceive a direct benefit. For these reasons, reliability performance should be tested under real-life conditions with the same conditions, providers, and levels of training as would be expected when the scale or item is in actual use. This has been conducted only for a limited number of the scales or items discussed above and has very rarely been tested for combined assessment tools.

A particularly relevant concern for implementing a common assessment tool is the need to achieve reliability across care settings. Different health settings naturally vary in their approaches to patients and care. Providers in different settings may have different levels of training, different experience, and a different framework for viewing patients. Many of the provider-tested measures of physical function would require significant training to achieve agreement,[230] particularly if the type of provider performing the assessment will vary across settings. Even short interview items can be strongly influenced by the interviewer's approach to the patient.[231] To some extent, having items collected at the hospital prior to discharge may decrease some of the variation that would be expected across settings and increase the chance for a more common assessment approach. This provides a strong argument in support of the current effort to implement a common assessment prior to hospital discharge. Even in a single setting, it remains important to ascertain whether the type of clinician collecting the data matters, and if so, specify who should provide which elements.

2. How Items will be Collected (Modes of Administration)

In a related issue for evaluating older adults, variations in mode of administration present another important factor that must be considered in efforts to develop a care transition assessment. Mode of administration (patient, provider, or proxy response) is a significant source of variation for interview items, even within a given setting. Careful attention will need to be paid to determine the best information source for a given item because different sources (patient interview, proxy interview, provider interview or observation, medical record) do not always agree.[157, 232-247] Whether a patient's or a provider's views are preferred is important to consider when determining whether an item should be patient vs. provider response. In addition, research comparisons have even shown significant differences for the same items between self-administered and interviewer-administered questionnaires.[248] An algorithm identifying a priority order for information source may be necessary for some items when a primary source is not available. For example, a non-communicative patient may not be able to answer depression items, requiring other approaches for assessment.

3. Burden

While each of the above domains is individually important for improving outcomes, care transitions and care planning, it is necessary to consider multiple domains in order to fully capture a patient's health state.[85, 249] This is particularly challenging when assessing vulnerable older adults with multiple conditions and syndromes.[250, 251] In addition, the number of items that are proposed to measure some domains can be significant. For example, most IADL/ADL lists used in surveys contain 10-15 items, and many item responses require follow-up questions (e.g., identifying task modification, identifying need for human assistance). Thus, a full set of IADL/ADL items can be cumbersome to incorporate in surveys and could pose significant provider and respondent burden. Further adding to potential burden, many items and scales are subject to floor and ceiling effects. The potential for these effects can be somewhat minimized in more homogeneous groups such as long-stay nursing home residents. However, for more diverse populations or for assessing across population subgroups with potentially wide ranging health states, this can present a significant but not insurmountable challenge to achieving parsimony. Providers are less likely to see an item as burdensome if they see relevance and utility for care processes that they value.

Different data sources offer different advantages and disadvantages for balancing reduced burden against relevance and accuracy. These considerations will vary across domains and items. This is a particularly important issue for follow-up and post-discharge assessments. Collecting information from pre-existing administrative data might decrease provider and respondent burden. However, administrative data vary in quality and content across organizations, and has well recognized limitations for monitoring or assessing populations.[252-254] As noted above, many important conditions in this population are under-recognized and underreported, particularly in administrative data. For other items, such as ethnicity and race, administrative data vary in accuracy for particular racial groups, being plagued by frequent missing data and being more accurate for white and African American than for other groups.[255-261] Most importantly, administrative data do not capture functional status.

A simple telephone survey or questionnaire could make the assessment transportable across settings and organizations and potentially more representative of a patient's real health.[262] However, surveys can be time-consuming to administer and may be plagued by non-response. Not all domains have identified items that have been collected by telephone interviews. [263]

Cognitive impairment presents a particular measurement challenge for some members of this population. Most research shows that even persons with moderate cognitive impairment are able to self-report symptoms of depression and pain[264-270] as well as answer multidimensional assessments,[271] although this finding is not consistent across all item sets or across all studies.[268, 269, 272, 273] Other approaches are clearly needed for assessing persons who are non-communicative or with severe cognitive impairment. Despite this challenge, it remains important to assess symptoms in persons with cognitive impairment as they appear to be at increased risk for under-treatment.[274, 275]

4. Acceptability

Provider acceptance is important for encouraging accurate assessment and for incorporation of assessment findings into better care.[276] In addition to time and data collection burden discussed above, the subject matter of a particular item may limit acceptability. This appears to be particularly relevant for many of the SES variables. Providers express ethical concerns about considering race in clinical care,[73, 277] and correctly point to genetic heterogeneity that exists within socially constructed racial categories.[73] Many patients feel uncomfortable with being asked about race, ethnicity, or SES, although this might be mitigated if they understand that the query is meant to ensure equal quality.[278] Many SES variables present unique challenges for self-report and administrative data with potentially complex interactions and variations in conclusions depending on definitions applied.[279] Although less dramatic, other items may not be asked because of limited acceptability. Despite the widespread use of the GDS to screen for depression in research studies and geriatric training programs, it has not been incorporated into standard care, with some providers objecting to routine screening terminology such as "worthless." For both GDS and other tools with scales, providers may fail to understand the need to ask and respondents may resist answering several items that seem redundant, merely because they load onto a common factor and permit scaling.

5. Item Wording and Application to Subsets of the Population

In describing the domains above, we noted that some, such as gender, race, language, and cognitive status are associated with variations in proclivity to report and in ability to respond to items. In addition, for these groups, individual item responses for subgroups of respondents may vary in terms of how they relate to the underlying construct being measured. Analytic methods are available to assess and address this problem if the same set of items has been collected in the relevant subgroups. For example, concerns have been raised that the self-reported difficulty of IADL/ADL items may vary by gender because of traditional differences in role-function, [99, 280, 281] and the hierarchical order of IADL items has been shown to vary between age subgroups.[105] Item response theory (IRT) methods have been used to show that the relationship of these items to the construct of disability do not vary for women compared to men or for oldest old compared to younger old.[282] Systematic cross-cultural testing is lacking for many items that may be subject to cultural variations in performance.[99, 283] If data on items were available in sufficiently large and diverse samples, similar approaches could be used to examine other items across relevant population subgroups.

Specific item wording is important in the performance of a domain across population subgroups and overtime.[105, 228] The prevalence of self-reported difficulty and self-reported need-for-help varies by age,[227] and item wording changes the relationships of the items to the underlying construct of disability.[282] For example, the response categories used to assess IADL/ADL limitations can include difficulty vs. no difficulty, able vs. not able, receive-help vs. not receive-help, and does vs. does not do. Difficulty is not a simple mid-point value on a scale extending from no difficulty to receive-help.[226] The choice of item wording and response categories can lead to different prevalence and outcome estimates across studies and in the same population. [104, 284 226, 284-286] In addition, individual IADL/ADL items differ in the extent to which they vary across response categories.[104, 227] These variations affect the interpretation of IADL/ADL items and can make comparison of ordered lists difficult.

Item wording also circumscribes the conclusions that can be drawn from a given item. For example, the MCBS asks if the respondent received help, not if the respondent needed help. This may lead to differential identification of IADL/ADL problems across groups with different access to resources. For example, assistive devices can avert the need for assistance and unmet need for these devices may hasten dependency. In one study of a dementia population, persons with baseline unmet need were more likely to be placed in nursing homes, die or be lost to follow-up, even after considering IADL/ADL and cognitive status.[287]

Finally, and perhaps most importantly, it is not clear whether any type of assessment, in the absence of other systematic quality improvement activities, will improve outcomes.[288] It is increasingly recognized that assessing and noting these factors is a necessary but not sufficient condition for improving care and care outcomes. Indeed, the U.S. Preventive Services Task Force recommends screening adults for depression ***if*** the clinical practice has systems in place to assure accurate diagnosis, effective treatment, and follow-up.[112] In one research study, a comprehensive assessment initiated in the hospital and followed up for completion at home identified problems and led to some changes in therapy;[289] however, the effect was limited and six-month outcomes were not improved.[290] Coupling assessment with management can make a difference. In a controlled trial of Geriatric Evaluation and Management (GEM) Units that

combined assessment with clinical management, patients cared for in inpatient GEM units had less functional decline at 6 months than those receiving standard care, and patients in outpatient GEM care had better mental health outcomes.[291]

D. EXAMPLES OF AVAILABLE INSTRUMENTS AND SURVEYS FOR COMMUNITY-DWELLING POPULATIONS

1. Health Outcomes Survey (Medicare HOS, formerly known as the Health of Seniors Survey)

The HOS is part of the Health Employer Data Set (HEDIS®) collection of measures published by the National Committee for Quality Assurance (NCQA). Plans collect this data as part of NCQA accreditation (if sought). The 90-item HOS includes 3 components: the SF-36, items for case-mix risk adjustment, and items added by CMS for monitoring Medicare Managed Care Plans. Both a baseline random sample of 1,000 Medicare Beneficiaries in each plan (if enrollment population <1000, then all beneficiaries are surveyed) and two-year follow-up survey are obtained. Since 1997, CMS has required all Medicare Managed Care Organizations (MCOs) to collect and report HEDIS® measures and MCOs have submitted HOS to CMS since 1998.[292]

The HOS is derivative of SF-36, a core set of 36 items.[88, 90, 293, 294] A slightly different scoring system has been put forth for the proprietary MOS SF-36® (International Resource Center for Health Care Assessment, Boston MA) and the publicly available RAND 36-item Health Survey 1.0. Two point changes on the overall score have been associated with clinically meaningful changes in status.

The SF-36 includes eight scales that measure the following domains: physical function (10 items); physical role limitations (4 items); emotional role limitation (3 items), social role limitation (2 items); emotional well-being (5 items); pain (2 items); energy or fatigue (4 items); general health perceptions (5 items). It has repeatedly been shown to evidence strong psychometric properties. The internal reliability of the scales has been reported as: physical function (0.93); physical role limitations (0.84); emotional role limitation (0.83), social role limitation (0.85); emotional well-being (0.90); pain (0.82), energy or fatigue (0.87), general health perceptions (0.78). Two component scores can also be generated: a physical component and a mental component.

The SF-36 has been repeatedly shown to be a valid measure of health status in a wide range of patient populations,[88, 295-302] has been used as a gold-standard outcome measure in multiple studies, has shown sensitivity to change, and has shown convergent validity with differences in care quality and other health indexes or physiologic measures, including studies in vulnerable and frail populations.[291, 303-314] The MHI-5 has shown validity when compared to gold-standard measures for identifying major depression in functionally impaired elders (sensitivity = 78.7% and specificity = 72.1%),[315] and the mental health subscale is highly correlated with the Geriatric Depression Scale (GDS) in nursing home residents capable of self-report.[316] Some scales such as role-emotion showed less convergent validity in nursing home populations.[316] It has also been applied as a validation gold standard in multiple studies developing new health status and quality of life measure.[317-320] The SF-36 has been tested in persons with mild to moderate cognitive impairment and shown to have continued acceptable psychometric properties.[271]

The HOS is designed to be a self-administered, mailed survey. The SF-36 and HOS can also be interviewer administered. A Spanish language version of the HOS is available. The SF-36 has been translated and validated in multiple languages.[321-330]

The SF-36 may not capture all relevant components of quality of life for all patient subgroups.[331-334] It has been criticized as more relevant to younger populations and subject to floor effects.[335-337] One study of the SF-36 (without ADL/IADL items) in an acute stroke population found internal consistency of the scales and sensitivity to change over time, but poor correlation with accepted measures of stroke severity.[338] Likewise, abbreviated versions of the SF-36, such as the SF-20, have been more subject to floor effects in more severely ill populations.[315, 326] However, the HOS added more ADL items, and some analyses in more vulnerable elders reveal continued validity of scales.[339] A modified version of the SF-36 that expands dichotomous response options has been developed for veteran populations with a resulting reduction in ceiling and floor effects for the expanded item responses.[340]

2. Continuity of Care Record (CCR)

The Continuity of Care Record (CCR) aims to provide a structure and standard to facilitate documentation and transfer of health information across providers and settings. It is intended as a longitudinal snapshot of patient data. The structure aims to accommodate multiple clinical domains with required sections addressing conditions (active, resolved, admission and discharge diagnoses), allergies (pharmacy, dietary, general), and medications (history, administered, discharge, current). Optional sections allow capture of advance directives, functional status, procedures, encounters, family history, social history, immunizations, vital signs, lab results and plan of care. In order to enhance interoperability, the CCR does not allow individual users of the standard to add additional sections. The CCR standard process is set up to allow additional sections to be added through a formal ballot process.

There are two sets of constraints or rules on the standard. The CCR intends to apply HL-7 messaging standards and if used in an electronic format requires adherence to XML schema, which says what names tags must have and in which order. It also defines what is optional and what is required. Work is underway to develop a HL7 CDA implementation guide for the CCR, which would easily allow HL7 V3 messaging. The second set of constraints comes from the implementation guide. It contains the constraints that are not easily definable in a XML schema, such as which controlled vocabulary to use. The CCR intends to have the ability to capture coded data such as type of clinical document/encounter, conditions, allergies, laboratory, vital signs, radiology procedures, operative procedures, other procedures such as LP, aspiration – in LOINC or any other controlled vocabulary. CCR will allow for the coding of a particular data element in multiple controlled vocabularies. For example, the problem list could be coded through a controlled vocabulary such as ICD9, SNOMED, or ICPC. The final specification will also be influenced by current efforts to establish which controlled vocabularies can quickly be absorbed by users and software developers and by current efforts to work with controlled vocabulary developers to achieve complete coverage of the concepts that might be captured by the CCR.

CCR does not specify a particular standardized item, scale or question wording, aiming instead to identify controlled vocabulary that would allow capture and summary of medical record

content in a uniform presentation. Enumerations of items were part of the early development of the standard, but the consensus process required that they be removed. Work is ongoing to identify or develop standardized, controlled vocabularies for several of the optional sections such as advance directives, functional status, family history, and social history. The CCR Acceleration Task Force is also looking at enumerations for concepts such as <Status>, which are not covered by a controlled vocabulary. Some of the vendors in the CCR Acceleration Task Force have taken the initial enumeration work and have worked to refine it for their needs. The need for these enumerations is due to the fact that some domains within the CCR are not covered by current "free" controlled vocabularies (or any controlled vocabulary). CCR Acceleration Task Force Leaders have noted that no demographic elements have been specified in the standard.

There is no consensus as to which medication controlled vocabulary should be used. The two most widely available, NDC and RxNorm, have recognized but different limitations. Some believe that NDC is not modeled in a way that facilitates physician use or decision support. RxNorm may be modeled in a more accessible way, but is not complete.

The CCR has not been implemented. The generation of a longitudinal health summary that relies on controlled vocabularies and can enhance information exchange in an EHR holds great promise for improving care transfers. However, a lack of standardized survey or assessment items raises questions about the ability of a tool such as a CCR to track quality issues or outcomes across providers and patients. One could imagine that some process measures based on appropriate medications for diagnoses could be generated.

3. VA GEC Referral Form

This assessment form was adapted from the Michigan Choice Instrument which includes items related the MDS-HC. Its primary purpose is to improve care planning at discharge, including the identification of the appropriate level of care and support needs. It includes 91 questions across the following 22 sections: (1) referral source; (2) patient's living situation; (3) caregiver characteristics (contact information, relationship to the patient, type of support, access, ability to increase help); (4) language; (5) homebound status; (6) IADLs (meal preparation, housework, shopping, transportation, telephone use, medication management, managing finances), and recent change in IADL; (7) services in the home; (8) other (recent change in living arrangements, want other living situation; amount of physical activity, left alone, surrogate decision maker, advance directive); (9) skilled care and treatments (DME, special diets, tube feeding, IV, catheter, dialysis, wound care, pressure ulcer care and stage, frequency of nurse observation, rehabilitation, substance abuse treatment); (10) ADLs (bathing, dressing, eating, using the toilet; moving in bed, transfers, moving around indoors, mobility with wheelchair) and recent change in ADLs; (11) continence (urine, stool); (12) skin; (13) behaviors and symptoms (wandering, verbally abusive behaviors, physically abusive behaviors, resisting care, hallucinations); (14) cognitive status (ease of decision making, ability to make self understood, agitation or disorientation poses danger); (15) prognosis (recent flare, direct care staff perception of possibility for increased independence, limited life expectancy); (16) weight bearing status; (17) diet; (18) equipment needs; (19) supplies needed; (20) goals of care (check all that apply); (21) program to which referred; and (22) estimated duration of services.

The VA instrument includes many of the domains of interest for the uniform assessment instrument; however, little information is available on the scales that are being proposed in terms of precision at different levels of impairment, reliability, and sensitivity to change. The background materials that support the VA GEC referral form indicate that, in field trials, if "a provider who knows the patient" completes the form, the completion time is under 10 minutes. A study is currently underway in the VA to evaluate baseline referral practices prior to full implementation of the VA GEC, which is scheduled for early 2006.

4. Core Outcome and Comprehensive Assessment – Basic (COCOA-B)

The 174 items in the COCOA-B address multiple domains relevant to patient outcomes, care transitions and resource planning. It serves as a comprehensive health assessment reviewing the domains listed above as well as additional symptoms, sensory impairments, prognosis, and nutritional status. Sections and items include: demographics & SES (gender, age, Medicaid eligibility, ethnicity, race, marital status, primary language and English fluency, education); diagnoses (record up to 8, ICD-9, severity rating 0-4; acute or chronic); overall prognosis and life expectancy; pain (presence, severity, frequency, interfere with daily activities, intractable); pressure ulcers (presence, number at each stage, stage of most problematic, status of most problematic); select health behaviors; flu immunization status; vision; hearing; height and weight; hydration (oral fluid intake, skin tugor); nutrition risk (10 items); other symptoms (dyspnea, edema); continence (UI frequency, UI occurence, treatment for UTI, bowel incontence); falls; medication management (independence, adherence); adherence to therapy/interventions; self-rated health; 9 IADL/ADL items (difficulty, receive help); functional performance (lifting 10 lbs, walk ¼ mile); attendance at day health center; type of residence; with whom the patient lives; caregiver (existence, number, frequency of assistance, type of assistance); advance directives; anxiety, stress/concerns; 10 symptoms of reported or disturbed mood; frequency of 6 behavior problems; cognitive items (cognitive function, memory, judgment, ability to understand others, ability to express thoughts); satisfaction and problems with care and access; social isolation; self-rated quality of life; caregiver stress, coping and satisfaction with services and access; items completed by therapist include patient endurance, ability to perform 8 ADLs (5-point response scale) and 6 IADLs, prognosis for rehab and structural variables.

5. The Vulnerable Elders 13 Item Survey (VES-13)

The VES-13 was developed as a parsimonious approach to identifying older adults at significant risk for health decline in order to allow better targeting of evaluations and care.[11, 341] It was developed as part of the Assessing Care of Vulnerable Elders (ACOVE) project. Its 13 items address age, physical function, self-rated health, and IADL/ADLs. It is being used in several health care systems to screen for and identify elders warranting referral for more comprehensive geriatric assessment. It provides information on a set of elders for whom quality of care measurably improves 15-month outcomes.[342]

The VES-13 is provided as an example of the type of tool that can be developed when common factors and outcomes are available for a group of patients. Its development considered a range of domains important for health outcomes in older adults. IRT methods were employed to test the applicability of IADL/ADL items to population subgroups (male, female, oldest-old, younger-

old) and to test change in performance if response choices were varied. Because it was planned as an easily transportable paper and pencil tool, best subset analyses were used to decrease item burden within a domain and regression models were used to eliminate domains from the tool. Finally, an integer-based scoring approach was developed and tested in order to facilitate provider and patient assessment. It can be self-administered, administered by non-clinical office staff over the phone,[166] or in face-to-face interviews in less than 5 minutes (average telephone = 2.8 minutes).

Its brevity prevents it from capturing all of the functional status items that might be important for a given patient, screening for syndromes such as mood or pain, and from capturing all of the data that would be important for an individual's care transition. Because diagnoses did not add to its targeting ability, they are not included, but are clearly relevant for care transitions. It is subject to ceiling effects in healthy populations. However, its physical function and IADL/ADL items are highly correlated with mood, pain and cognition; reported problems with VES-13 items are being used to identify those patients warranting more detailed queries of these conditions and syndromes. Self-rated health also correlates with multiple dimensions of health, including disease diagnoses and function,[32, 343] and provides additional information even when these items are considered.[51, 344, 345]

E. CONCLUSION

Several domains are recognized as important for standardized, quality assessment of older adults in hospital and community settings. To select items for measuring domains in diverse populations, several challenges will need to be addressed. Some instruments exist that aim to improve assessment but none seem to offer the brevity, flexibility, adequate testing across provider and population subgroups, and a range that avoids ceiling and floor effects. The VA GEC and COCOA-B assess multiple domains, but feasibility and acceptability for providers remains to be determined. The CCR includes multiple domains, but has not been tested and fails to identify specific item sets for measuring those domains. While this achieves the desired flexibility for an electronic health record, it would prohibit comparisons across providers, an important objective for a national data collection effort. For many domains, it would also limit its utility and ease of translation for different providers because they may not be familiar with or understand how to translate particular measures applied by other providers. The SF-36 has the widest experience and offers the advantage of current use as part of HEDIS. Domains like medications that are untapped by the SF-36 could be added.

F. REFERENCES

1. Saliba D, Waite MS. Risk Assessment and Identification. In: Mezey M, Mottrel M, et al, eds. The Encyclopedia of Elder Care: The Comprehensive Resource on Geriatric and Social Care. New York: Srpinger Publishing Company; 2000.

2. Stuck AE, Walthert JM, Nikolaus T, Bèula CJ, Hohmann C, Beck JC. Risk factors for functional status decline in community-living elderly people: a systematic literature review. Soc Sci Med. Feb 1999;48(4):445-469.

3. Li AK, Covinsky KE, Sands LP, Fortinsky RH, Counsell SR, Landefeld CS. Reports of financial disability predict functional decline and death in older patients discharged from the hospital. J Gen Intern Med. Feb 2005;20(2):168-174.

4. Melzer D, Izmirlian G, Leveille SG, Guralnik JM. Educational differences in the prevalence of mobility disability in old age: the dynamics of incidence, mortality, and recovery. J Gerontol B Psychol Sci Soc Sci. Sep 2001;56(5):S294-301.

5. McCally M, Haines A, Fein O, Addington W, Lawrence RS, Cassel CK. Poverty and ill health: physicians can, and should, make a difference. Ann Intern Med. Nov 1 1998;129(9):726-733.

6. Lantz PM, House JS, Lepkowski JM, Williams DR, Mero RP, Chen J. Socioeconomic factors, health behaviors, and mortality: results from a nationally representative prospective study of US adults. Jama. Jun 3 1998;279(21):1703-1708.

7. Camacho TC, Strawbridge WJ, Cohen RD, Kaplan GA. Functional Ability in the Oldest Old: Cumulative Impact of Risk Factors from the Preceding Two Decades. Journal of Aging and Health. 1993;5(4):439-454.

8. Goldman N, Korenman S, Weinstein R. Marital status and health among the elderly. Soc Sci Med. Jun 1995;40(12):1717-1730.

9. Crimmins EM, Saito Y. Trends in healthy life expectancy in the United States, 1970-1990: gender, racial, and educational differences. Soc Sci Med. Jun 2001;52(11):1629-1641.

10. Palmore EB, Nowlin JB, Wang HS. Predictors of function among the old-old: a 10-year follow-up. J Gerontol. Mar 1985;40(2):244-250.

11. Saliba D, Elliott M, Rubenstein LZ, et al. The Vulnerable Elders Survey: a tool for identifying vulnerable older people in the community. J Am Geriatr Soc. Dec 2001;49(12):1691-1699.

12. Inouye SK, Peduzzi PN, Robison JT, Hughes JS, Horwitz RI, Concato J. Importance of functional measures in predicting mortality among older hospitalized patients. Jama. Apr 15 1998;279(15):1187-1193.

13. Salive ME, Satterfield S, Ostfeld AM, Wallace RB, Havlik RJ. Disability and cognitive impairment are risk factors for pneumonia-related mortality in older adults. Public Health Rep. May-Jun 1993;108(3):314-322.

14. Ferrucci L, Guralnik JM, Baroni A, Tesi G, Antonini E, Marchionni N. Value of combined assessment of physical health and functional status in community-dwelling aged: a prospective study in Florence, Italy. J Gerontol. Mar 1991;46(2):M52-56.

15. Gill TM, Williams CS, Richardson ED, Tinetti ME. Impairments in physical performance and cognitive status as predisposing factors for functional dependence among nondisabled older persons. J Gerontol A Biol Sci Med Sci. Nov 1996;51(6):M283-288.

16. Pincus T, Brooks RH, Callahan LF. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. Ann Intern Med. Jan 1 1994;120(1):26-34.

17. Fried LP, Kronmal RA, Newman AB, et al. Risk factors for 5-year mortality in older adults: the Cardiovascular Health Study. Jama. Feb 25 1998;279(8):585-592.

18. Sarkisian CA, Liu H, Gutierrez PR, Seeley DG, Cummings SR, Mangione CM. Modifiable risk factors predict functional decline among older women: a prospectively validated clinical prediction tool. The Study of Osteoporotic Fractures Research Group. J Am Geriatr Soc. Feb 2000;48(2):170-178.

19. Kaplan GA, Strawbridge WJ, Camacho T, Cohen RD. Factors Associated With Change in Physical Functioning in the Elderly: A Six-Year Prospective Study. Journal of Aging and Health. 1993;5(1):140-153.

20. Cho CY, Alessi CA, Cho M, et al. The association between chronic illness and functional change among participants in a comprehensive geriatric assessment program. J Am Geriatr Soc. Jun 1998;46(6):677-682.

21. Cornoni-Huntley JC, Foley DJ, Guralnik JM. Co-morbidity analysis: a strategy for understanding mortality, disability and use of health care facilities of older people. Int J Epidemiol. 1991;20:S8-17.

22. Boult C, Altmann M, Gilbertson D, Yu C, Kane RL. Decreasing disability in the 21st century: the future effects of controlling six fatal and nonfatal conditions. Am J Public Health. Oct 1996;86(10):1388-1393.

23. Furner SE, Rudberg MA, Cassel CK. Medical conditions differentially affect the development of IADL disability: implications for medical care and research. Gerontologist. Aug 1995;35(4):444-450.

24. Mor V, Wilcox V, Rakowski W, Hiris J. Functional transitions among the elderly: patterns, predictors, and related hospital use. Am J Public Health. Aug 1994;84(8):1274-1280.

25. Gill TM, Robison JT, Tinetti ME. Predictors of recovery in activities of daily living among disabled older persons living in the community. J Gen Intern Med. Dec 1997;12(12):757-762.

26. Crimmins EM, Saito Y. Getting Better and Getting Worse: Transitions in Functional Status among Older Americans. Journal of Aging and Health. 1993;5(1):3-36.

27. Manton KG, Corder LS, Stallard E. Estimates of change in chronic disability and institutional incidence and prevalence rates in the U.S. elderly population from the 1982, 1984, and 1989 National Long Term Care Survey. J Gerontol. Jul 1993;48(4):S153-166.

28. Beckett LA, Brock DB, Lemke JH, et al. Analysis of change in self-reported physical function among older persons in four population studies. Am J Epidemiol. Apr 15 1996;143(8):766-778.

29. Branch LG, Katz S, Kniepmann K, Papsidero JA. A prospective study of functional status among community elders. Am J Public Health. Mar 1984;74(3):266-268.

30. Jette AM, Branch LG, Berlin J. Musculoskeletal impairments and physical disablement among the aged. J Gerontol. Nov 1990;45(6):M203-208.

31. Rakowski W, Fleishman JA, Mor V, Bryant SA. Self-Assessments of Health and Mortality Among Older Persons. Res Aging. 1993;15:92-116.

32. Bernard SL, Kincade JE, Konrad TR, et al. Predicting mortality from community surveys of older adults: the importance of self-rated functional ability. J Gerontol B Psychol Sci Soc Sci. May 1997;52(3):S155-163.

33. Borawski EA, Kinney JM, Kahana E. The meaning of older adults' health appraisals: congruence with health status and determinant of mortality. J Gerontol B Psychol Sci Soc Sci. May 1996;51(3):S157-170.

34. Idler EL, Kasl SV, Lemke JH. Self-evaluated health and mortality among the elderly in New Haven, Connecticut, and Iowa and Washington counties, Iowa, 1982-1986. Am J Epidemiol. Jan 1990;131(1):91-103.

35. Lawrence RH, Jette AM. Disentangling the disablement process. J Gerontol B Psychol Sci Soc Sci. Jul 1996;51(4):S173-182.

36. Seeman TE, Berkman LF, Kohout F, Lacroix A, Glynn R, Blazer D. Intercommunity variations in the association between social ties and mortality in the elderly. A comparative analysis of three communities. Ann Epidemiol. Jul 1993;3(4):325-335.

37. Boult C, Dowd B, McCaffrey D, Boult L, Hernandez R, Krulewitch H. Screening elders for risk of hospital admission. J Am Geriatr Soc. Aug 1993;41(8):811-817.

38. Coleman EA, Wagner EH, Grothaus LC, Hecht J, Savarino J, Buchner DM. Predicting hospitalization and functional decline in older health plan enrollees: are administrative data as accurate as self-report? J Am Geriatr Soc. Apr 1998;46(4):419-425.

39. Brody KK, Johnson RE, Douglas Ried L. Evaluation of a self-report screening instrument to predict frailty outcomes in aging populations. Gerontologist. Apr 1997;37(2):182-191.

40. Kaplan RC, Tirschwell DL, Longstreth WT, Jr., et al. Vascular events, mortality, and preventive therapy following ischemic stroke in the elderly. Neurology. Sep 27 2005;65(6):835-842.

41. Majumdar SR, Gurwitz JH, Soumerai SB. Undertreatment of hyperlipidemia in the secondary prevention of coronary artery disease. J Gen Intern Med. Dec 1999;14(12):711-717.

42. McLaughlin TJ, Soumerai SB, Willison DJ, et al. Adherence to national guidelines for drug treatment of suspected acute myocardial infarction: evidence for undertreatment in women and the elderly. Arch Intern Med. Apr 8 1996;156(7):799-805.

43. Benoit SR, Mendelsohn AB, Nourjah P, Staffa JA, Graham DJ. Risk factors for prolonged QTc among US adults: Third National Health and Nutrition Examination Survey. Eur J Cardiovasc Prev Rehabil. Aug 2005;12(4):363-368.

44. Ness J, Aronow WS, Newkirk E, McDanel D. Underutilization of calcium and vitamin D by older adults in a large general internal medicine practice. Am J Ther. Mar-Apr 2005;12(2):113-116.

45. Schall MB, Flannery J. Undertreatment of women with heart failure: a reversible outcome on hospital readmission. Lippincotts Case Manag. Nov-Dec 2004;9(6):250-253.

46. Green CR, Wheeler JR, LaPorte F. Clinical decision making in pain management: Contributions of physician and patient characteristics to variations in practice. J Pain. Feb 2003;4(1):29-39.

47. Kroenke K, Spitzer RL. Gender differences in the reporting of physical and somatoform symptoms. Psychosom Med. Mar-Apr 1998;60(2):150-155.

48. Lai H, Lai S, Krongrad A, Trapido E, Page JB, McCoy CB. The effect of marital status on survival in late-stage cancer patients: an analysis based on. Int J Behav Med. 1999;6(2):150-176.

49. Gore JL, Kwan L, Saigal CS, Litwin MS. Marriage and mortality in bladder carcinoma. Cancer. Sep 15 2005;104(6):1188-1194.

50. Grande GE, Addington-Hall JM, Todd CJ. Place of death and access to home care services: are certain patient groups at a disadvantage? Soc Sci Med. Sep 1998;47(5):565-579.

51. Rakowski W, Mor V. The association of physical activity with mortality among older adults in the Longitudinal Study of Aging (1984-1988). J Gerontol. Jul 1992;47(4):M122-129.

52. Cox C. Outcomes of hospitalization: factors influencing the discharges of African American and white dementia patients. Soc Work Health Care. 1996;23(1):23-38.

53. Jette AM, Tennstedt S, Crawford S. How does formal and informal community care affect nursing home use? J Gerontol B Psychol Sci Soc Sci. Jan 1995;50(1):S4-S12.

54. Ayanian JZ, Zaslavsky AM, Guadagnoli E, et al. Patients' perceptions of quality of care for colorectal cancer by race, ethnicity, and language. J Clin Oncol. Sep 20 2005;23(27):6576-6586.

55. Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. Jama. May 12 1993;269(18):2386-2391.

56. Kaplan RC, Bhalodkar NC, Brown DL, White J, Brown EJ Jr. Differences by Age and Race/Ethnicity in Knowledge about Hypercholesterolemia. Cariol Rev. Jan/Feb 2006;14(1):1-6.

57. O'Neill C, Harrington C, Kitchener M, Saliba D. Quality of care in nursing homes: an analysis of relationships among profit, quality, and ownership. Med Care. Dec 2003;41(12):1318-1330.

58. Farley JF, Cline RR, Gupta K. Racial variations in antiresorptive medication use: results from the 2000 Medical Expenditure Panel Survey (MEPS). Osteoporosis International. Nov 29 2005.

59. Morris AM, Billingsley KG, Baxter NN, Baldwin LM. Racial disparities in rectal cancer treatment: a population-based analysis. Arch Surg. Feb 2004;139(2):151-155; discussion 156.

60. Melzer D, Lan TY, Tom BD, Deeg DJ, Guralnik JM. Variation in thresholds for reporting mobility disability between national population subgroups and studies. J Gerontol A Biol Sci Med Sci. Dec 2004;59(12):1295-1303.

61. Joo JH, Solano FX, Mulsant BH, Reynolds CF, Lenze EJ. Predictors of adequacy of depression management in the primary care setting. Psychiatr Serv. Dec 2005;46(12):1524-1528.

62. Adams AS, Zhang F, Mah C, et al. Race differences in long-term diabetes management in an HMO. Diabetes Care. Dec 2005;28(12):2844-2849.

63. Okonofua EC, Cutler NE, Lackland DT, Egan BM. Ethnic differences in older americans: awareness, knowledge, and beliefs about hypertension. Am J Hypertens. Jul 2005;18(7):972-979.

64. Gansler T, Henley SJ, Stein K, Nehl EJ, Smigal C, Slaughter E. Sociodemographic determinants of cancer treatment health literacy. Cancer. Aug 1 2005;104(3):653-660.

65. Rehman SU, Hutchison FN, Hendrix K, Okonofua EC, Egan BM. Ethnic differences in blood pressure control among men at Veterans Affairs clinics and other health care sites. Arch Intern Med. May 9 2005;165(9):1041-1047.

66. Golightly YM, Dominick KL. Racial variations in self-reported osteoarthritis symptom severity among veterans. Aging Clin Exp Res. Aug 2005;17(4):264-269.

67. Nau DP, Ellis JJ, Kline-Rogers EM, Mallya U, Eagle KA, Erickson SR. Gender and perceived severity of cardiac disease: evidence that women are "tougher". Am J Med. Nov 2005;118(11):1256-1261.

68. McDermott MM, Guralnik JM, Ferrucci L, et al. Functional decline in lower-extremity peripheral arterial disease: Associations with comorbidity, gender, and race. J Vasc Surg. Dec 2005;42(6):1131-1137.

69. Riehle JF, Lackland DT, Okonofua EC, Hendrix KH, Egan BM. Ethnic differences in the treatment and control of hypertension in patients with diabetes. J Clin Hypertens (Greenwich). Aug 2005;7(8):445-454.

70. Tracy JK, Meyer WA, Flores RH, Wilson PD, Hochberg MC. Racial differences in rate of decline in bone mass in older men: the Baltimore men's osteoporosis study. J Bone Miner Res. Jul 2005;20(7):1228-1234.

71. BiDil for heart failure. Med Lett Drugs Ther. Sep 26 2005;47(1218):77-78.

72. Taylor AL, Ziesche S, Yancy C, et al. Combination of isosorbide dinitrate and hydralazine in blacks with heart failure. N Engl J Med. Nov 11 2004;351(20):2049-2057.

73. Bloche MG. Race-based therapeutics. N Engl J Med. Nov 11 2004;351(20):2035-2037.

74. Cunningham WE, Hays RD, Duan N, et al. The effect of socioeconomic status on the survival of people receiving care for HIV infection in the United States. Journal of Health Care for the Poor and Underserved. Nov 2005;16(4):655-676.

75. Wolf MS, Gazmararian JA, Baker DW. Health literacy and functional health status among older adults. Arch Intern Med. Sep 26 2005;165(17):1946-1952.

76. Praska JL, Kripalani S, Seright AL, Jacobson TA. Identifying and assisting low-literacy patients with medication use: a survey of community pharmacies. Ann Pharmacother. Sep 2005;39(9):1441-1445.

77. Verbrugge LM, Jette AM. The disablement process. Soc Sci Med. Jan 1994;38(1):1-14.

78. Katz S, Stroud MW, 3rd. Functional assessment in geriatrics. A review of progress and directions. J Am Geriatr Soc. Mar 1989;37(3):267-271.

79. Wenger NS, Shekelle PG. Assessing care of vulnerable elders: ACOVE project overview. Ann Intern Med. Oct 16 2001;135(8):642-646.

80. Nagi SZ. An epidemiology of disability among adults in the United States. Milbank Mem Fund Q Health Soc. Fall 1976;54(4):439-467.

81. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Engl J Med. Mar 2 1995;332(9):556-561.

82. Giampaoli S, Ferrucci L, Cecchi F, et al. Hand-grip strength predicts incident disability in non-disabled older men. Age Ageing. May 1999;28(3):283-288.

83. Winograd CH, Lindenberger EC, Chavez CM, Mauricio MP, Shi H, Bloch DA. Identifying hospitalized older patients at varying risk for physical performance decline: a new approach. J Am Geriatr Soc. May 1997;45(5):604-609.

84. Judge JO, Schechtman K, Cress E. The relationship between physical performance measures and independence in instrumental activities of daily living. The FICSIT Group. Frailty and Injury: Cooperative Studies of Intervention Trials. J Am Geriatr Soc. Nov 1996;44(11):1332-1341.

85. Reuben DB, Siu AL, Kimpau S. The predictive validity of self-report and performance-based measures of function and health. J Gerontol. Jul 1992;47(4):M106-110.

86. Hays RD, Marshall GN, Wang EY, Sherbourne CD. Four-year cross-lagged associations between physical and mental health in the Medical Outcomes Study. J Consult Clin Psychol. Jun 1994;62(3):441-449.

87. Onder G, Penninx BW, Ferrucci L, Fried LP, Guralnik JM, Pahor M. Measures of physical performance and risk for progressive and catastrophic disability: results from the Women's Health and Aging Study. J Gerontol A Biol Sci Med Sci. Jan 2005;60(1):74-79.

88. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. Jun 1992;30(6):473-483.

89. Rozzini R, Frisoni GB, Bianchetti A, Zanetti O, Trabucchi M. Physical Performance Test and Activities of Daily Living scales in the assessment of health status in elderly people. J Am Geriatr Soc. Oct 1993;41(10):1109-1113.

90. Stewart AL, Sherbourne C, Hays RD, et al. Summary and discussion of MOS measures. In: Steward AL, Ware JE, eds. Measuring functioning and well-being: The Medical Outcomes Study Approach. Durham: Duke University Press; 1992:345-371.

91. Sayers SP, Guralnik JM, Thombs LA, Fielding RA. Effect of leg muscle contraction velocity on functional performance in older men and women. J Am Geriatr Soc. Mar 2005;53(3):467-471.

92. Duncan PW, Studenski S, Chandler J, Prescott B. Functional reach: predictive validity in a sample of elderly male veterans. J Gerontol. May 1992;47(3):M93-98.

93. Reuben DB, Siu AL. An objective measure of physical function of elderly outpatients. The Physical Performance Test. J Am Geriatr Soc. Oct 1990;38(10):1105-1112.

94. Binder EF, Miller JP, Ball LJ. Development of a test of physical performance for the nursing home setting. Gerontologist. Oct 2001;41(5):671-679.

95. Kruiansky J, Gurland B. The performance test of activities of daily living. Int J Aging Hum Dev. 1976;7(4):343-352.

96. Ferrucci L, Guralnik JM, Bandeen-Roche KJ, Lafferty ME, Pahor M, Fried LP. Physical Performance Measures Report accessed on January 2, 2006 at http://www.grc.nia.nih.gov/branches/ledb/whasbook/chap4/chap4.htm.

97. MULTIDISCIPLINARY studies of illness in aged persons. II. A new classification of functional status in activities of daily living. J Chronic Dis. Jan 1959;9(1):55-62.

98. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies Of Illness In The Aged. The Index Of Adl: A Standardized Measure Of Biological And Psychosocial Function. Jama. Sep 21 1963;185:914-919.

99. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist. Autumn 1969;9(3):179-186.

100. Roy C, Togneri J, Hay E, B P. An inter-rater reliability study of the Barthel Index. International Journal of Rehabilitation Research. 1988;11:67-70.

101. LaPlante MP. The demographics of disability. Milbank Q. 1991;69:55-77.

102. Plehn K, Marcopulos BA, McLain CA. The relationship between neuropsychological test performance, social functioning, and instrumental activities of daily living in a sample of rural older adults. Clin Neuropsychol. Feb 2004;18(1):101-113.

103. Siu AL, Reuben DB, Hays RD. Hierarchical measures of physical function in ambulatory geriatrics. J Am Geriatr Soc. Oct 1990;38(10):1113-1119.

104. Myers AM. The clinical Swiss army knife. Empirical evidence on the validity of IADL functional status measures. Med Care. May 1992;30(5):MS96-111.

105. Fillenbaum GG. Screening the elderly. A brief instrumental activities of daily living measure. J Am Geriatr Soc. Oct 1985;33(10):698-706.

106. Spector WD, Katz S, Murphy JB, Fulton JP. The hierarchical relationship between activities of daily living and instrumental activities of daily living. J Chronic Dis. 1987;40(6):481-489.

107. Katz S, Ford AB. Prognosis after strokes. Part II. Long-term course of 159 patients. Medicine (Baltimore). May 1966;45(3):236-246.

108. Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. Gerontologist. Spring 1970;10(1):20-30.

109. Davis RB, Iezzoni LI, Phillips RS, Reiley P, Coffman GA, Safran C. Predicting in-hospital mortality. The importance of functional status information. Med Care. Sep 1995;33(9):906-921.

110. Rozzini R, Sabatini T, Cassinadri A, et al. Relationship between functional loss before hospital admission and mortality in elderly persons with medical illness. J Gerontol A Biol Sci Med Sci. Sep 2005;60(9):1180-1183.

111. Reuben DB, Rubenstein LV, Hirsch SH, Hays RD. Value of functional status as a predictor of mortality: results of a prospective study. Am J Med. Dec 1992;93(6):663-669.

112. (USPSTF) USPSTF. Guide to Clinical Preventive Services. Agency for Healthcare Research and Quality [website]. November 2003; Available at: http://www.ahrq.gov/clinic/uspstfix.htm. Accessed January 16, 2006.

113. Pringle KE, Ahern FM, Heller DA, Gold CH, Brown TV. Potential for alcohol and prescription drug interactions in older people. J Am Geriatr Soc. Nov 2005;53(11):1930-1936.

114. Moore AA, Endo JO, Carter MK. Is there a relationship between excessive drinking and functional impairment in older persons? J Am Geriatr Soc. Jan 2003;51(1):44-49.

115. Fink A, Tsai MC, Hays RD, et al. Comparing the alcohol-related problems survey (ARPS) to traditional alcohol screening measures in elderly outpatients. Arch Gerontol Geriatr. Feb 2002;34(1):55-78.

116. Simonsick EM, Lafferty ME, Phillips CL, et al. Risk due to inactivity in physically capable older adults. Am J Public Health. Oct 1993;83(10):1443-1450.

117. Teno JM. Looking beyond the "form" to complex interventions needed to improve end-of-life care. J Am Geriatr Soc. Sep 1998;46(9):1170-1171.

118. Meyers JL, Moore C, McGrory A, Sparr J, Ahern M. Physician orders for life-sustaining treatment form: honoring end-of-life directives for nursing home residents. J Gerontol Nurs. Sep 2004;30(9):37-46.

119. Schmidt TA, Hickman SE, Tolle SW, Brooks HS. The Physician Orders for Life-Sustaining Treatment program: Oregon emergency medical technicians' practical experiences and attitudes. J Am Geriatr Soc. Sep 2004;52(9):1430-1434.

120. Hickman SE, Tolle SW, Brummel-Smith K, Carley MM. Use of the Physician Orders for Life-Sustaining Treatment program in Oregon nursing facilities: beyond resuscitation status. J Am Geriatr Soc. Sep 2004;52(9):1424-1429.

121. Cantor MD. Improving advance care planning: lessons from POLST. Physician Orders for Life-Sustaining Treatment. J Am Geriatr Soc. Oct 2000;48(10):1343-1344.

122. Tolle SW, Tilden VP, Nelson CA, Dunn PM. A prospective study of the efficacy of the physician order form for life-sustaining treatment. J Am Geriatr Soc. Sep 1998;46(9):1097-1102.

123. Agèuero-Torres H, Fratiglioni L, Guo Z, Viitanen M, von Strauss E, Winblad B. Dementia is the major cause of functional dependence in the elderly: 3-year follow-up data from a population-based study. Am J Public Health. Oct 1998;88(10):1452-1456.

124. Sauvaget C, Yamada M, Fujiwara S, Sasaki H, Mimori Y. Dementia as a predictor of functional disability: a four-year follow-up study. Gerontology. Jul-Aug 2002;48(4):226-233.

125. Chodosh J, Petitti DB, Elliott M, et al. Physician recognition of cognitive impairment: evaluating the need for improvement. J Am Geriatr Soc. Jul 2004;52(7):1051-1059.

126. Greiner PA, Snowdon DA, Schmitt FA. The loss of independence in activities of daily living: the role of low normal cognitive function in elderly nuns. Am J Public Health. Jan 1996;86(1):62-66.

127. Boult C, Kane RL, Louis TA, Boult L, McCaffrey D. Chronic conditions that lead to functional limitation in the elderly. J Gerontol. Jan 1994;49(1):M28-36.

128. Mor V, Murphy J, Masterson-Allen S, et al. Risk of functional decline among well elders. J Clin Epidemiol. 1989;42(9):895-904.

129. Pettersen R, Dahl T, Wyller TB. Prediction of long-term functional outcome after stroke rehabilitation. Clin Rehabil. Mar 2002;16(2):149-159.

130. Bruce ML, Seeman TE, Merrill SS, Blazer DG. The impact of depressive symptomatology on physical disability: MacArthur Studies of Successful Aging. Am J Public Health. Nov 1994;84(11):1796-1799.

131. Unèutzer J, Katon W, Callahan CM, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. Jama. Dec 11 2002;288(22):2836-2845.

132. Corti MC, Guralnik JM, Salive ME, Sorkin JD. Serum albumin level and physical disability as predictors of mortality in older persons. Jama. Oct 5 1994;272(13):1036-1042.

133. Boult C, Murphy J, Sloane P, Mor V, Drone C. The relation of dizziness to functional decline. J Am Geriatr Soc. Sep 1991;39(9):858-861.

134. Otto CM, Lind BK, Kitzman DW, Gersh BJ, Siscovick DS. Association of aortic-valve sclerosis with cardiovascular mortality and morbidity in the elderly. N Engl J Med. Jul 15 1999;341(3):142-147.

135. Dargie HJ, McMurray JJ, McDonagh TA. Heart failure--implications of the true size of the problem. J Intern Med. Apr 1996;239(4):309-315.

136. Strawbridge WJ, Wallhagen MI, Shema SJ, Kaplan GA. Negative consequences of hearing impairment in old age: a longitudinal analysis. Gerontologist. Jun 2000;40(3):320-326.

137. Wolinsky FD, Fitzgerald JF, Stump TE. The effect of hip fracture on mortality, hospitalization, and functional status: a prospective study. Am J Public Health. Mar 1997;87(3):398-403.

138. Incalzi AR, Capparella O, Gemma A, et al. A simple method of recognizing geriatric patients at risk for death and disability. J Am Geriatr Soc. Jan 1992;40(1):34-38.

139. Laird RD, Studenski S, Perera S, Wallace D. Fall history is an independent predictor of adverse health outcomes and utilization in the elderly. Am J Manag Care. Dec 2001;7(12):1133-1138.

140. Bischoff-Ferrari HA, Lingard EA, Losina E, et al. Psychosocial and geriatric correlates of functional status after total hip replacement. Arthritis Rheum. Oct 15 2004;51(5):829-835.

141. Iezzoni LI. Using administrative data to study persons with disabilities. Milbank Q. 2002;80(2):347-379.

142. Soumerai SB, McLaughlin TJ, Spiegelman D, Hertzmark E, Thibault G, Goldman L. Adverse outcomes of underuse of beta-blockers in elderly survivors of acute myocardial infarction. Jama. Jan 8 1997;277(2):115-121.

143. Department of Health and Human Services CfMMS. Medicare Prescription Drug Data Strategy: Improving Evidence for Patient Care Through the Medicare Prescription Drug Benefit. Pharmaceutical Care Management Association [website] Available at: http://www.pcmanet.org/events/presentation/CMS%20Paper%20-%20Data%20Strategy%20for%20Medicare%20Drug%20Benefit%20Overview.pdf. Accessed Jan 6, 2006.

144. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol. Apr 2005;58(4):323-337.

145. Coffin CS, Saunders C, Thomas CM, Loewen AH, Ghali WA, Campbell NR. Validity of ICD-9-CM administrative data for determining eligibility for pneumococcal vaccination triggers. Am J Med Qual. May-Jun 2005;20(3):158-163.

146. Lee DS, Donovan L, Austin PC, et al. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. Med Care. Feb 2005;43(2):182-188.

147. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. J Clin Epidemiol. Feb 2004;57(2):131-141.

148. Szeto HC, Coleman RK, Gholami P, Hoffman BB, Goldstein MK. Accuracy of computerized outpatient diagnoses in a Veterans Affairs general medicine clinic. Am J Manag Care. Jan 2002;8(1):37-43.

149. Singh JA, Holmgren AR, Noorbaloochi S. Accuracy of Veterans Administration databases for a diagnosis of rheumatoid arthritis. Arthritis Rheum. Dec 15 2004;51(6):952-957.

150. Harrold LR, Yood RA, Straus W, et al. Challenges of estimating health service utilization for osteoarthritis patients on a population level. J Rheumatol. Sep 2002;29(9):1931-1936.

151. MacLean CH, Louie R, Venus PJ, et al. Comparison of Administrative Data and Medical Records to Measure the Quality of Medical Care Provided to Vulnerable Older Patients. Med Care. In press.

152. Berlowitz DR, Brand HK, Perkins C. Geriatric syndromes as outcome measures of hospital care: can administrative data be used? J Am Geriatr Soc. Jun 1999;47(6):692-696.

153. Lillquist PP. Challenges in surveillance of dementias in New York State. Prev Chronic Dis. Jan 2004;1(1):A08.

154. Joray S, Wietlisbach V, Bèula CJ. Cognitive impairment in elderly medical inpatients: detection and associated six-month outcomes. Am J Geriatr Psychiatry. Nov-Dec 2004;12(6):639-647.

155. Nierenberg AA. Current perspectives on the diagnosis and treatment of major depressive disorder. Am J Manag Care. Sep 2001;7(11):S353-366.

156. Hirschfeld RM, Keller MB, Panico S, et al. The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. Jama. Jan 22-29 1997;277(4):333-340.

157. Schmaling KB, Hernandez DV. Detection of depression among low-income Mexican Americans in primary care. J Health Care Poor Underserved. Nov 2005;16(4):780-790.

158. Weiner D, Peterson B, Ladd K, McConnell E, Keefe F. Pain in nursing home residents: an exploration of prevalence, staff perspectives, and practical aspects of measurement. Clin J Pain. Jun 1999;15(2):92-101.

159. Wu N, Miller SC, Lapane K, Roy J, Mor V. The quality of the quality indicator of pain derived from the minimum data set. Health Serv Res. Aug 2005;40(4):1197-1216.

160. McNeill JA, Sherwood GD, Starck PL. The hidden error of mismanaged pain: a systems approach. J Pain Symptom Manage. Jul 2004;28(1):47-58.

161. Inouye SK, Leo-Summers L, Zhang Y, Bogardus ST, Jr., Leslie DL, Agostini JV. A chart-based method for identification of delirium: validation compared with interviewer ratings using the confusion assessment method. J Am Geriatr Soc. Feb 2005;53(2):312-318.

162. Rubenstein LZ, Solomon DH, Roth CP, et al. Detection and management of falls and instability in vulnerable elders by community physicians. J Am Geriatr Soc. Sep 2004;52(9):1527-1531.

163. Sloss EM, Solomon DH, Shekelle PG, et al. Selecting target conditions for quality of care improvement in vulnerable older adults. J Am Geriatr Soc. Apr 2000;48(4):363-369.

164. Seymour RM, Routledge PA. Important drug-drug interactions in the elderly. Drugs Aging. Jun 1998;12(6):485-494.

165. Pot AM, Deeg DJ, Twisk JW, Beekman AT, Zarit SH. The longitudinal relationship between the use of long-term care and depressive symptoms in older adults. Gerontologist. Jun 2005;45(3):359-369.

166. Wenger NS, Solomon DH, Roth CP, et al. The quality of medical care provided to vulnerable community-dwelling older patients. Ann Intern Med. Nov 4 2003;139(9):740-747.

167. Sheikh JI, Yesavage JA. Geriatric Depression Scale: recent evidence and development of a shorter version. Clin Gerontol. 1986;5:165-172.

168. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res. 1982-83 1982;17(1):37-49.

169. Radloff LS. The Center for Epidemilogical Studies--Depression Scale: A self-report depression scale for research in the general population. Applied Psychological Measurements. 1977;3:385-401.

170. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. Jama. Nov 10 1999;282(18):1737-1744.

171. Spitzer RL, Williams JB, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. Jama. Dec 14 1994;272(22):1749-1756.

172. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression. Two questions are as good as many. J Gen Intern Med. Jul 1997;12(7):439-445.

173. Lèowe B, Kroenke K, Herzog W, Grèafe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). J Affect Disord. Jul 2004;81(1):61-66.

174. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. Sep 2001;16(9):606-613.

175. Areâan PA, Ayalon L, Hunkeler E, et al. Improving depression care for older, minority patients in primary care. Med Care. Apr 2005;43(4):381-390.

176. Williams JBW. A Structured Interview Guide for the Hamilton Depression Rating Scale. Archives of General Psychiatry. 1998;45:742-747.

177. Baker FM, Miller CL. Screening a skilled nursing home population for depression. J Geriatr Psychiatry Neurol. Oct-Dec 1991;4(4):218-221.

178. Lyness JM, Noel TK, Cox C, King DA, Conwell Y, Caine ED. Screening for depression in elderly primary care patients. A comparison of the Center for Epidemiologic Studies-Depression Scale and the Geriatric Depression Scale. Arch Intern Med. Feb 24 1997;157(4):449-454.

179. Gerety MB, Williams JW, Jr., Mulrow CD, et al. Performance of case-finding tools for depression in the nursing home: influence of clinical and functional characteristics and selection of optimal threshold scores. J Am Geriatr Soc. Oct 1994;42(10):1103-1109.

180. Kilbourne AM, Schulberg HC, Post EP, Rollman BL, Belnap BH, Pincus HA. Translating evidence-based depression management services to community-based primary care practices. Milbank Q. 2004;82(4):631-659.

181. Williams JW, Noel PH, Cordes JA, Ramirez G, Pignone M. Is this patient clinically depressed? Journal of the American Medical Association. 2002;287:1160-1170.

182. Katon WJ, Schoenbaum M, Fan MY, et al. Cost-effectiveness of improving primary care treatment of late-life depression. Arch Gen Psychiatry. Dec 2005;62(12):1313-1320.

183. Callahan CM, Kroenke K, Counsell SR, et al. Treatment of depression improves physical functioning in older adults. J Am Geriatr Soc. Mar 2005;53(3):367-373.

184. Lin EH, Katon W, Von Korff M, et al. Effect of improving depression care on pain and functional outcomes among older adults with arthritis: a randomized controlled trial. Jama. Nov 12 2003;290(18):2428-2429.

185. Mahoney J, Drinka TJ, Abler R, et al. Screening for depression: single question versus GDS. J Am Geriatr Soc. Sep 1994;42(9):1006-1008.

186. Blank K, Gruman C, Robison JT. Case-finding for depression in elderly people: balancing ease of administration with validity in varied treatment settings. J Gerontol A Biol Sci Med Sci. Apr 2004;59(4):378-384.

187. Boyle VL, Roychoudhury C, Beniak R, Cohn L, Bayer A, Katz I. Recognition and management of depression in skilled-nursing and long-term care settings: evolving targets for quality improvement. Am J Geriatr Psychiatry. May-Jun 2004;12(3):288-295.

188. The management of persistent pain in older persons. J Am Geriatr Soc. Jun 2002;50(6):S205-224.

189. Pain: The Fifth Vital Sign™. American Pain Society [website] Available at: http://www.ampainsoc.org/advocacy/fifth.htm. Accessed Jan 17, 2006.

190. Pain as the 5th Vital Sign: Take 5. VA Trainee Pocket Card [website] Available at: http://www.va.gov/oaa/pocketcard/pain.asp. Accessed Jan 6, 2006.

191. Pain Assessment: The Fifth Vital Sign. State of California Department of Consumer Affairs [website]. July 2000; Available at: http://www.rn.ca.gov/practice/pdf/npr-b-27.pdf. Accessed Jan 6, 2006.

192. Taylor LJ, Harris J, Epps CD, Herr K. Psychometric evaluation of selected pain intensity scales for use with cognitively impaired and cognitively intact older adults. Rehabil Nurs. Mar-Apr 2005;30(2):55-61.

193. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. Pain. Sep 1975;1(3):277-299.

194. Ferrell BA, Stein WM, Beck JC. The Geriatric Pain Measure: validity, reliability and factor analysis. J Am Geriatr Soc. Dec 2000;48(12):1669-1673.

195. Campbell SE, Seymour DG, Primrose WR. A systematic literature review of factors affecting outcome in older medical patients admitted to hospital. Age Ageing. Mar 2004;33(2):110-115.

196. Lenze EJ, Munin MC, Dew MA, et al. Adverse effects of depression and cognitive impairment on rehabilitation participation and recovery from hip fracture. Int J Geriatr Psychiatry. May 2004;19(5):472-478.

197. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. Nov 1975;12(3):189-198.

198. Libon DJ, Swenson RA, Barnoski EJ, Sands LP. Clock drawing as an assessment tool for dementia. Arch Clin Neuropsychol. Oct 1993;8(5):405-415.

199. Lavery LL, Starenchak SM, Flynn WB, Stoeff MA, Schaffner R, Newman AB. The clock drawing test is an independent predictor of incident use of 24-hour care in a retirement community. J Gerontol A Biol Sci Med Sci. Jul 2005;60(7):928-932.

200. Wong A, Mok VC, Yim P, et al. The executive clock drawing task (CLOX) is a poor screening test for executive dysfunction in Chinese elderly patients with subcortical ischemic vascular disease. J Clin Neurosci. Jun 2004;11(5):493-497.

201. Diegelman NM, Gilbertson AD, Moore JL, Banou E, Meager MR. Validity of the Clock Drawing Test in predicting reports of driving problems in the elderly. BMC Geriatr. Oct 29 2004;4:10.

202. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The mini-cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. Int J Geriatr Psychiatry. Nov 2000;15(11):1021-1027.

203. Scanlan J, Borson S. The Mini-Cog: receiver operating characteristics with expert and naèive raters. Int J Geriatr Psychiatry. Feb 2001;16(2):216-222.

204. Borson S, Scanlan JM, Chen P, Ganguli M. The Mini-Cog as a screen for dementia: validation in a population-based sample. J Am Geriatr Soc. Oct 2003;51(10):1451-1454.

205. Callahan CM, Unverzagt FW, Hui SL, Perkins AJ, Hendrie HC. Six-item screener to identify cognitive impairment among potential subjects for clinical research. Med Care. Sep 2002;40(9):771-781.

206. Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. Am J Psychiatry. Jun 1983;140(6):734-739.

207. Kahn RL, Goldfarb AI, Pollack M, Peck A. Brief objective measures for the determination of mental status in the aged. Am J Psychiatry. Oct 1960;117:326-328.

208. de Jager CA, Budge MM, Clarke R. Utility of TICS-M for the assessment of cognitive function in older adults. Int J Geriatr Psychiatry. Apr 2003;18(4):318-324.

209. Crooks VC, Clark L, Petitti DB, Chui H, Chiu V. Validation of multi-stage telephone-based identification of cognitive impairment and dementia. BMC Neurol. Apr 2005;5(1):8.

210. Ely EW, Siegel MD, Inouye SK. Delirium in the intensive care unit: an under-recognized syndrome of organ dysfunction. Semin Respir Crit Care Med. 2001;22(2):115-126.

211. Leslie DL, Zhang Y, Holford TR, Bogardus ST, Leo-Summers LS, Inouye SK. Premature death associated with delirium at 1-year follow-up. Arch Intern Med. Jul 25 2005;165(14):1657-1662.

212. Marcantonio ER, Kiely DK, Simon SE, et al. Outcomes of older people admitted to postacute facilities with delirium. J Am Geriatr Soc. Jun 2005;53(6):963-969.

213. Marcantonio ER, Flacker JM, Michaels M, Resnick NM. Delirium is independently associated with poor functional recovery after hip fracture. J Am Geriatr Soc. Jun 2000;48(6):618-624.

214. Inouye SK, Schlesinger MJ, Lydon TJ. Delirium: a symptom of how hospital care is failing older persons and a window to improve quality of hospital care. Am J Med. May 1999;106(5):565-573.

215. Fick DM, Kolanowski AM, Waller JL, Inouye SK. Delirium superimposed on dementia in a community-dwelling managed care population: a 3-year retrospective study of occurrence, costs, and utilization. J Gerontol A Biol Sci Med Sci. Jun 2005;60(6):748-753.

216. Leslie DL, Zhang Y, Bogardus ST, Holford TR, Leo-Summers LS, Inouye SK. Consequences of preventing delirium in hospitalized older adults on nursing home costs. J Am Geriatr Soc. Mar 2005;53(3):405-409.

217. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med. Dec 15 1990;113(12):941-948.

218. Inouye SK, Foreman MD, Mion LC, Katz KH, Cooney LM, Jr. Nurses' recognition of delirium and its symptoms: comparison of nurse and researcher ratings. Arch Intern Med. Nov 12 2001;161(20):2467-2473.

219. American Geriatrics Society, British Geriatrics Society, and American Academy of Orthopaedic Surgeons Panel on Falls Prevention. Guideline for the Prevention of Falls in Older Persons. American Geriatrics Society [website]. 2001; Available at: http://www.americangeriatrics.org/products/positionpapers/Falls.pdf. Accessed Jan 6, 2006.

220. Fletcher PC, Hirdes JP. Restriction in activity associated with fear of falling among community-based seniors using home care services. Age Ageing. May 2004;33(3):273-279.

221. Murphy SL, Dubin JA, Gill TM. The development of fear of falling among community-living older women: predisposing factors and subsequent fall events. J Gerontol A Biol Sci Med Sci. Oct 2003;58(10):M943-947.

222. Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls. A prospective study. Jama. May 12 1989;261(18):2663-2668.

223. Tinetti ME, Williams CS. The effect of falls and fall injuries on functioning in community-dwelling older persons. J Gerontol A Biol Sci Med Sci. Mar 1998;53(2):M112-119.

224. Wilson MM, Miller DK, Andresen EM, Malmstrom TK, Miller JP, Wolinsky FD. Fear of falling and related activity restriction among middle-aged African Americans. J Gerontol A Biol Sci Med Sci. Mar 2005;60(3):355-360.

225. J²rstad EC, Hauer K, Becker C, Lamb SE. Measuring the psychological outcomes of falling: a systematic review. J Am Geriatr Soc. Mar 2005;53(3):501-510.

226. Wiener JM, Hanley RJ, Clark R, Van Nostrand JF. Measuring the activities of daily living: comparisons across national surveys. J Gerontol. Nov 1990;45(6):S229-237.

227. Jette AM. How measurement techniques influence estimates of disability in older populations. Soc Sci Med. Apr 1994;38(7):937-942.

228. Roberts G. Age effects and health appraisal: a meta-analysis. J Gerontol B Psychol Sci Soc Sci. Jan 1999;54(1):S24-30.

229. Brook RH, Lohr KN. Efficacy, effectiveness, variations, and quality. Boundary-crossing research. Med Care. May 1985;23(5):710-722.

230. Jette AM, Jette DU, Ng J, Plotkin DJ, Bach MA. Are performance-based measures sufficiently reliable for use in multicenter trials? Musculoskeletal Impairment (MSI) Study Group. J Gerontol A Biol Sci Med Sci. Jan 1999;54(1):M3-6.

231. Kane RL, Bershadsky B, Kane RA, et al. Using resident reports of quality of life to distinguish among nursing homes. Gerontologist. Oct 2004;44(5):624-632.

232. Rothman ML, Hedrick SC, Bulcroft KA, Hickam DH, Rubenstein LZ. The validity of proxy-generated scores as measures of patient health status. Med Care. Feb 1991;29(2):115-124.

233. Hall JA, Epstein AM, McNeil BJ. Multidimensionality of health status in an elderly population. Construct validity of a measurement battery. Med Care. Mar 1989;27(3):S168-177.

234. Conroy MB, Majchrzak NE, Silverman CB, et al. Measuring provider adherence to tobacco treatment guidelines: a comparison of electronic medical record review, patient survey, and provider survey. Nicotine Tob Res. Apr 2005;7:S35-43.

235. Subramanian U, Weiner M, Gradus-Pizlo I, Wu J, Tu W, Murray MD. Patient perception and provider assessment of severity of heart failure as predictors of hospitalization. Heart Lung. Mar-Apr 2005;34(2):89-98.

236. Boyer F, Novella JL, Morrone I, Jolly D, Blanchard F. Agreement between dementia patient report and proxy reports using the Nottingham Health Profile. Int J Geriatr Psychiatry. Nov 2004;19(11):1026-1034.

237. von Essen L. Proxy ratings of patient quality of life--factors related to patient-proxy agreement. Acta Oncol. 2004;43(3):229-234.

238. Lorence D. Regional variation in medical classification agreement: benchmarking the coding gap. J Med Syst. Oct 2003;27(5):435-443.

239. Justice AC, Rabeneck L, Hays RD, Wu AW, Bozzette SA. Sensitivity, specificity, reliability, and clinical validity of provider-reported symptoms: a comparison with self-reported symptoms. Outcomes Committee of the AIDS Clinical Trials Group. J Acquir Immune Defic Syndr. Jun 1 1999;21(2):126-133.

240. Pierre U, Wood-Dauphinee S, Korner-Bitensky N, Gayton D, Hanley J. Proxy use of the Canadian SF-36 in rating health status of the disabled elderly. J Clin Epidemiol. Nov 1998;51(11):983-990.

241. Sainfort F, Becker M, Diamond R. Judgments of quality of life of individuals with severe mental disorders: Patient self-report versus provider perspectives. Am J Psychiatry. Apr 1996;153(4):497-502.

242. Berlowitz DR, Du W, Kazis L, Lewis S. Health-related quality of life of nursing home residents: differences in patient and provider perceptions. J Am Geriatr Soc. Jul 1995;43(7):799-802.

243. Grossman SA, Sheidler VR, Swedeen K, Mucenski J, Piantadosi S. Correlation of patient and caregiver ratings of cancer pain. J Pain Symptom Manage. Feb 1991;6(2):53-57.

244. Yasuda N, Zimmerman S, Hawkes WG, Gruber-Baldini AL, Hebel JR, Magaziner J. Concordance of proxy-perceived change and measured change in multiple domains of function in older persons. J Am Geriatr Soc. Jul 2004;52(7):1157-1162.

245. Rogers JC, Holm MB, Beach S, et al. Concordance of four methods of disability assessment using performance in the home as the criterion method. Arthritis Rheum. Oct 15 2003;49(5):640-647.

246. Santos-Eggimann B, Zobel F, Bâerod AC. Functional status of elderly home care users: do subjects, informal and professional caregivers agree? J Clin Epidemiol. Mar 1999;52(3):181-186.

247. Magaziner J, Zimmerman SI, Gruber-Baldini AL, Hebel JR, Fox KM. Proxy reporting in five areas of functional status. Comparison with self-reports and observations of performance. Am J Epidemiol. Sep 1 1997;146(5):418-428.

248. Reuben DB, Valle LA, Hays RD, Siu AL. Measuring physical function in community-dwelling older persons: a comparison of self-administered, interviewer-administered, and performance-based measures. J Am Geriatr Soc. Jan 1995;43(1):17-23.

249. Siu AL, Reuben DB, Ouslander JG, Osterweil D. Using multidimensional health measures in older persons to identify risk of hospitalization and skilled nursing placement. Qual Life Res. Aug 1993;2(4):253-261.

250. Tinetti ME, Bogardus ST, Jr., Agostini JV. Potential pitfalls of disease-specific guidelines for patients with multiple conditions. N Engl J Med. Dec 30 2004;351(27):2870-2874.

251. McCormick WC, Boling PA. Multimorbidity and a comprehensive medicare care-coordination benefit. J Am Geriatr Soc. Dec 2005;53(12):2227-2228.

252. Iezzoni LI. The risks of risk adjustment. Jama. Nov 19 1997;278(19):1600-1607.

253. Iezzoni LI. Assessing quality using administrative data. Ann Intern Med. Oct 15 1997;127(8):666-674.

254. Bradley EH, Besdine R. Outcomes-based quality improvement: reducing the data collection burden. J Am Geriatr Soc. Apr 1998;46(4):534-535.

255. Harwell TS, Hansen D, Moore KR, Jeanotte D, Gohdes D, Helgerson SD. Accuracy of race coding on American Indian death certificates, Montana 1996-1998. Public Health Rep. Jan-Feb 2002;117(1):44-49.

256. Kelly JJ, Chu SY, Diaz T, Leary LS, Buehler JW. Race/ethnicity misclassification of persons reported with AIDS. The AIDS Mortality Project Group and The Supplement to HIV/AIDS Surveillance Project Group. Ethn Health. Mar 1996;1(1):87-94.

257. Bach PB, McClellan MB. A Prescription for a Modern Medicare Program. N Engl J Med. Dec 29 2005;353(26):2733-2744.

258. Data Issues Brief. VIRECC [website]. March 2004; Available at: http://www.virec.research.med.va.gov/References/DataIssuesBrief/DIB-0403.pdf. Accessed Jan 6, 2006.

259. Graber JM, Corkum BE, Sonnenfeld N, Kuehnert PL. Underestimation of cardiovascular disease mortality among Maine American Indians: the role of procedural and data errors. Am J Public Health. May 2005;95(5):827-830.

260. Blustein J. The reliability of racial classifications in hospital discharge abstract data. Am J Public Health. Jun 1994;84(6):1018-1021.

261. Patel DA, Knowles A, Schwartz AG, Schwartz K. Evaluation of African-American and white racial classification in a surveillance, epidemiology, and end results cancer registry. Ethn Dis. Autumn 2005;15(4):713-719.

262. National Research Council. Health Performance Measurement in the Public Sector: Principles and Policies for Implementing an Information Network.: National Academy Press; 1999.

263. Marcantonio ER, Michaels M, Resnick NM. Diagnosing delirium by telephone. J Gen Intern Med. Sep 1998;13(9):621-623.

264. McGivney SA, Mulvihill M, Taylor B. Validating the GDS depression screen in the nursing home. J Am Geriatr Soc. May 1994;42(5):490-492.

265. Sutcliffe C, Cordingley L, Burns A, et al. A new version of the geriatric depression scale for nursing and residential home populations: the geriatric depression scale (residential) (GDS-12R). Int Psychogeriatr. Jun 2000;12(2):173-181.

266. Snow AL, Kunik ME, Molinari VA, et al. Accuracy of self-reported depression in persons with dementia. J Am Geriatr Soc. Mar 2005;53(3):389-396.

267. Katz IR. Diagnosis and treatment of depression in patients with Alzheimer's disease and other dementias. J Clin Psychiatry. 1998;59:38-44.

268. Ferrell BA, Ferrell BR, Rivera L. Pain in cognitively impaired nursing home patients. J Pain Symptom Manage. Nov 1995;10(8):591-598.

269. Kaasalainen S, Crook J. A comparison of pain-assessment tools for use with elderly long-term-care residents. Can J Nurs Res. Dec 2003;35(4):58-71.

270. Manz BD, Mosier R, Nusser-Gerlach MA, Bergstrom N, Agrawal S. Pain assessment in the cognitively impaired and unimpaired elderly. Pain Manag Nurs. Dec 2000;1(4):106-115.

271. Novella JL, Jochum C, Ankri J, Morrone I, Jolly D, Blanchard F. Measuring general health status in dementia: practical and methodological issues in using the SF-36. Aging (Milano). Oct 2001;13(5):362-369.

272. Gilley DW, Wilson RS. Criterion-related validity of the Geriatric Depression Scale in Alzheimer's disease. J Clin Exp Neuropsychol. Aug 1997;19(4):489-499.

273. Kafonek S, Ettinger WH, Roca R, Kittner S, Taylor N, German PS. Instruments for screening for depression and dementia in a long-term care facility. Journal of the American Geriatrics Society. Jan 1989;37(1):29-34.

274. Allen RS, Thorn BE, Fisher SE, et al. Prescription and dosage of analgesic medication in relation to resident behaviors in the nursing home. J Am Geriatr Soc. Apr 2003;51(4):534-538.

275. Kaasalainen S, Middleton J, Knezacek S, et al. Pain and cognitive status in the institutionalized elderly: perceptions & interventions. J Gerontol Nurs. Aug 1998;24(8):24-31; quiz 50-21.

276. Rubenstein LV, Mittman BS, Yano EM, Mulrow CD. From understanding health care provider behavior to improving health care: the QUERI framework for quality improvement. Quality Enhancement Research Initiative. Med Care. Jun 2000;38(6):I129-141.

277. Bowser R. Race as a proxy for drug response: the dangers and challenges of ethnic drugs. De Paul Law Rev. Spring 2004;53(3):1111-1126.

278. Baker DW, Cameron KA, Feinglass J, et al. Patients' attitudes toward health care providers collecting information about their race and ethnicity. J Gen Intern Med. Oct 2005;20(10):895-900.

279. Braveman PA, Cubbin C, Egerter S, et al. Socioeconomic Status in Health Research: One Size Does Not Fit All. JAMA. Dec 2005;294(22):2879-2888.

280. Kempen GI, Miedema I, Ormel J, Molenaar W. The assessment of disability with the Groningen Activity Restriction Scale. Conceptual framework and psychometric properties. Soc Sci Med. Dec 1996;43(11):1601-1610.

281. Allen SM, Mor V, Raveis V, Houts P. Measurement of need for assistance with daily activities: quantifying the influence of gender roles. J Gerontol. Jul 1993;48(4):S204-211.

282. Saliba D, Orlando M, Wenger NS, Hays RD, Rubenstein LZ. Identifying a short functional disability screen for older persons. J Gerontol A Biol Sci Med Sci. Dec 2000;55(12):M750-756.

283. Alonso J, Black C, Norregaard JC, et al. Cross-cultural differences in the reporting of global functional capacity: an example in cataract patients. Med Care. Jun 1998;36(6):868-878.

284. Gill TM, Robison JT, Tinetti ME. Difficulty and dependence: two components of the disability continuum among community-living older persons. Ann Intern Med. Jan 15 1998;128(2):96-101.

285. Langlois JA, Maggi S, Harris T, et al. Self-report of difficulty in performing functional activities identifies a broad range of disability in old age. J Am Geriatr Soc. Dec 1996;44(12):1421-1428.

286. Fried LP, Herdman SJ, Kuhn KE, Rubin G, Turano K. Preclinical Disability - Hypotheses About the Bottom of the Iceburg. Journal of Aging and Health. 1991;3(2):285-300.

287. Gaugler JE, Kane RL, Kane RA, Newcomer R. Unmet care needs and key outcomes in dementia. Journal of the American Geriatrics Society. Dec 2005;53(12):2098-2105.

288. Simon GE, Fleck M, Lucas R, Bushnell DM. Prevalence and predictors of depression treatment in an international primary care study. Am J Psychiatry. Sep 2004;161(9):1626-1634.

289. Kravitz RL, Reuben DB, Davis JW, et al. Geriatric home assessment after hospital discharge. J Am Geriatr Soc. Dec 1994;42(12):1229-1234.

290. Siu AL, Kravitz RL, Keeler E, et al. Postdischarge geriatric assessment of hospitalized frail elderly patients. Arch Intern Med. Jan 8 1996;156(1):76-81.

291. Cohen HJ, Feussner JR, Weinberger M, et al. A controlled trial of inpatient and outpatient geriatric evaluation and management. N Engl J Med. Mar 21 2002;346(12):905-912.

292. Health Outcomes Survey Overview. Centers for Medicare and Medicaid Services [website]. Updated Dec 14, 2005; Available at: www.cms.hhs.gov/HOS/. Accessed Dec 30, 2005.

293. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. Health Econ. Oct 1993;2(3):217-227.

294. Brook RH, Ware JE, Jr., Davies-Avery A, et al. Overview of adult health measures fielded in Rand's health insurance study. Med Care. Jul 1979;17(7):iii-x, 1-131.

295. Haywood KL, Garratt AM, Fitzpatrick R. Quality of life in older people: a structured review of generic self-assessed health instruments. Qual Life Res. Sep 2005;14(7):1651-1668.

296. Butterworth P, Crosier T. The validity of the SF-36 in an Australian National Household Survey: demonstrating the applicability of the Household Income and Labour Dynamics in Australia (HILDA) Survey to examination of health inequalities. BMC Public Health. Oct 7 2004;4:44.

297. Ho AK, Robbins AO, Walters SJ, Kaptoge S, Sahakian BJ, Barker RA. Health-related quality of life in Huntington's disease: a comparison of two generic instruments, SF-36 and SIP. Mov Disord. Nov 2004;19(11):1341-1348.

298. Gâomez-Besteiro MI, Santiago-Pâerez MI, Alonso-Hernâandez A, Valdâes-Caänedo F, Rebollo-Alvarez P. Validity and reliability of the SF-36 questionnaire in patients on the waiting list for a kidney transplant and transplant patients. Am J Nephrol. May-Jun 2004;24(3):346-351.

299. Stansfeld SA, Roberts R, Foot SP. Assessing the validity of the SF-36 General Health Survey. Qual Life Res. Apr 1997;6(3):217-224.

300. Prieto L, Alonso J, Ferrer M, Antâo JM. Are results of the SF-36 health survey and the Nottingham Health Profile similar? A comparison in COPD patients. Quality of Life in COPD Study Group. J Clin Epidemiol. Apr 1997;50(4):463-473.

301. Riley ED, Bangsberg DR, Perry S, Clark RA, Moss AR, Wu AW. Reliability and validity of the SF-36 in HIV-infected homeless and marginally housed individuals. Qual Life Res. Dec 2003;12(8):1051-1058.

302. McHorney CA, Ware JE, Jr., Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care. Jan 1994;32(1):40-66.

303. Ware JE, Jr., Bayliss MS, Rogers WH, Kosinski M, Tarlov AR. Differences in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems. Results from the Medical Outcomes Study. Jama. Oct 2 1996;276(13):1039-1047.

304. Riazi A, Hobart JC, Lamping DL, Fitzpatrick R, Thompson AJ. Evidence-based measurement in multiple sclerosis: the psychometric properties of the physical and psychological dimensions of three quality of life rating scales. Mult Scler. Aug 2003;9(4):411-419.

305. Olney SJ, Nymark J, Brouwer B, et al. A Randomized Controlled Trial of Supervised Versus Unsupervised Exercise Programs for Ambulatory Stroke Survivors. Stroke. Jan 12 2006;37(1).

306. Boini S, Briancon S, Guillemin F, Galan P, Hercberg S. Occurrence of coronary artery disease has an adverse impact on health-related quality of life: a longitudinal controlled trial. Int J Cardiol. Dec 27 2005.

307. Pizzi LT, Weston CM, Goldfarb NI, et al. Impact of Chronic Conditions on Quality of Life in Patiens with Inflammatory Bowel Disease. Inflamm Bowel Dis. Jan 2006;12(1):47-52.

308. Bradshaw PJ, Jamrozik K, Gilfillan IS, Thompson PL. Return to work after coronary artery bypass surgery in a population of long-term survivors. Heart Lung Circ. Sep 2005;14(3):191-196.

309. Dogan E, Erkoc R, Eryonucu B, Saryarlioglu H, Agargun MY. Relation between depression, some laboratory parameters, and quality of life in hemodialysis patients. Ren Fail. 2005;27(6):695-699.

310. Cranney AB, Coyle D, Hopman WM, Hum V, Power B, Tugwell PS. Prospective evaluation of preferences and quality of life in women with hip fractures. J Rheumatol. Dec 2005;32(12):2393-2399.

311. Purser JL, Weinberger M, Cohen HJ, et al. Walking speed predicts health status and hospital costs for frail elderly male veterans. J Rehabil Res Dev. Jul-Aug 2005;42(4):535-546.

312. Lee AJ, Morgan CL, Conway P, Currie CJ. Characterisation and comparison of health-related quality of life for patients with renal failure. Curr Med Res Opin. Nov 2005;21(11):1777-1783.

313. Van den Berg JP, Kalmijn S, Lindeman E, et al. Multidisciplinary ALS care improves quality of life in patients with ALS. Neurology. Oct 25 2005;65(8):1264-1267.

314. Huang TT, Liang SH. A randomized clinical trial of the effectiveness of a discharge planning intervention in hospitalized elders with hip fracture due to falling. J Clin Nurs. Nov 2005;14(10):1193-1201.

315. Friedman B, Heisel M, Delavan R. Validity of the SF-36 Five-Item Mental Health Index for Major Depression in Functionally Impaired, Community-Dwelling Elderly Patients. J Am Geriatr Soc. Nov 2005;53(11):1978-1985.

316. Andresen EM, Gravitt GW, Aydelotte ME, Podgorski CA. Limitations of the SF-36 in a sample of nursing home residents. Age Ageing. Oct 1999;28(6):562-566.

317. Ferrer M, Cordoba J, Garin O, et al. Validity of the Spanish version of the Chronic Liver Disease Questionnaire (CLDQ) as a standard outcome for quality of life assesment. Liver Transpl. 2006 2006;12(1):95-104.

318. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. Gen Hosp Psychiatry. Nov 2006;28(1):968-983.

319. Fortin M, Hudon C, Dubois MF, Almirall J, Lapointe L, Soubhi H. Comparative assessment of three different indices of multimorbidity for studies of health-related quality of life. Health Qual Life Outcomes. Nov 2005.

320. Tan EK, Fook-Chong S, Lum SY, Thumboo J. Validation of a short disease specific quality of life scale for hemifacial spasm: correlation with SF-36. J Neurol Neurosurg Psychiatry. Dec 2005;76(12):1707-1710.

321. Anagnostopoulos F, Niakas D, Pappa E. Construct validation of the Greek SF-36 Health Survey. Qual Life Res. Oct 2005;14(8):1959-1965.

322. Kebede D, Alem A, Shibre T, Negash A, Deyassa N, Beyero T. Health related quality of life (SF-36) survey in Butajira, rural Ethiopia: normative data and evaluation of reliability and validity. Ethiop Med J. Oct 2004;42(4):289-297.

323. Yamazaki S, Fukuhara S, Green J. Usefulness of five-item and three-item Mental Health Inventories to screen for depressive symptoms in the general population of Japan. Health Qual Life Outcomes. Aug 8 2005;3:48.

324. Montazeri A, Goshtasebi A, Vahdaninia M, Gandek B. The Short Form Health Survey (SF-36): translation and validation study of the Iranian version. Qual Life Res. Apr 2005;14(3):875-882.

325. Rodrâiguez-Artalejo F, Guallar-Castillâon P, Pascual CR, et al. Health-related quality of life as a predictor of hospital readmission and death among patients with heart failure. Arch Intern Med. Jun 13 2005;165(11):1274-1279.

326. Peek MK, Ray L, Patel K, Stoebner-May D, Ottenbacher KJ. Reliability and validity of the SF-36 among older Mexican Americans. Gerontologist. Jun 2004;44(3):418-425.

327. Ayuso-Mateos JL, Lasa L, Vâazquez-Barquero JL, Oviedo A, Diez-Manrique JF. Measuring health status in psychiatric community surveys: internal and external validity of the Spanish version of the SF-36. Acta Psychiatr Scand. Jan 1999;99(1):26-32.

328. Alonso J, Prieto L, Ferrer M, et al. Testing the measurement properties of the Spanish version of the SF-36 Health Survey among male patients with chronic obstructive pulmonary disease. Quality of Life in COPD Study Group. J Clin Epidemiol. Nov 1998;51(11):1087-1094.

329. Pinar R. Reliability and construct validity of the SF-36 in Turkish cancer patients. Qual Life Res. Feb 2005;14(1):259-264.

330. Lubetkin EI, Jia H, Gold MR. Use of the SF-36 in low-income Chinese American primary care patients. Med Care. Apr 2003;41(4):447-457.

331. Post M, Noreau L. Quality of life after spinal cord injury. J Neurol Phys Ther. Sept 2005;29(3):139-146.

332. Scheingraber S, Scheingraber T, Brauckhoff M, Dralle H. Comparison between a general and a disease-specific health-related quality-of-life questionnaire in patients after pancreatic surgery. J Hepatobiliary Pancreat Surg. 2005;12(4):290-297.

333. Sararaks S, Azman AB, Low LL, et al. Validity and reliability of the SF-36: the Malaysian context. Med J Malaysia. Jun 2005;60(2):163-179.

334. Silveira E, Taft C, Sundh V, Waern M, Palsson S, Steen B. Performance of the SF-36 health survey in screening for depressive and anxiety disorders in an elderly female Swedish population. Qual Life Res. Jun 2005;14(5):1263-1274.

335. Hagen S, Bugge C, Alexander H. Psychometric properties of the SF-36 in the early post-stroke phase. J Adv Nurs. Dec 2003;44(5):461-468.

336. Quintana JM, Escobar A, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after hip joint replacement. Osteoarthritis Cartilage. Dec 2005;13(12):1076-1083.

337. Bourke SC, McColl E, Shaw PJ, Gibson GJ. Validation of quality of life instruments in ALS. Amyotroph Lateral Scler Other Motor Neuron Disord. Mar 2004;5(1):55-60.

338. Bindman AB, Keane D, Lurie N. Measuring health changes among severely ill patients. The floor phenomenon. Med Care. Dec 1990;28(12):1142-1152.

339. Kazis LE, Miller DR, Clark JA, et al. Improving the response choices on the veterans SF-36 health survey role functioning scales: results from the Veterans Health Study. J Ambul Care Manage. Jul-Sep 2004;27(3):263-280.

340. Hoeymans N, Feskens EJ, Kromhout D, van den Bos GA. The contribution of chronic conditions and disabilities to poor self-rated health in elderly men. J Gerontol A Biol Sci Med Sci. Oct 1999;54(10):M501-506.

341. Min L, Elliot M, Wenger NS, Saliba D. Higher VES-13 Scores Predict Death and Functional Decline at One Year Among Vulnerable Elders. Journal of the American Geriatrics Society. In press.

342. Higashi T, Shekelle PG, Adams JL, et al. Quality of care is associated with survival in vulnerable older patients. Ann Intern Med. Aug 16 2005;143(4):274-281.

343. Wolinsky FD, Tierney WM. Self-rated health and adverse health outcomes: an exploration and refinement of the trajectory hypothesis. J Gerontol B Psychol Sci Soc Sci. Nov 1998;53(6):S336-340.

344. Rakowski W, Mor V, Hiris J. An investigation of nonresponse to self-assessment of health by older persons. Associations with mortality. J Aging Health. Nov 1994;6(4):469-488.

345. Idler EL, Kasl S. Health perceptions and survival: do global evaluations of health status really predict mortality? J Gerontol. Mar 1991;46(2):S55-65.

Chapter 4
Hospital Discharge Assessment and Data Items that Facilitate Appropriate Placement and Efficient Care Transitions

A. INTRODUCTION

This chapter focuses on two of the three explicit primary purposes for a uniform assessment tool: the facilitation of decision making regarding appropriate PAC placement and the provision of core patient information to promote safe and efficient care transitions between care settings. The third primary purpose, assessment of outcomes across settings and over time, is addressed in Chapter 6.

In conceptualizing the approach to developing a Uniform Patient Assessment Instrument (UPAI) that could be used across acute and PAC settings, MedPAC recommended that the initial focus should be to develop an instrument that would be completed at the time of discharge from the acute hospital. This assessment would enable providers to adequately initiate PAC in the most appropriate setting and provide a baseline uniform characterization of patients for longitudinal monitoring as they transfer between acute and post-acute care, as well as between different PAC settings. In so doing, this tool would be integral to promoting greater beneficiary-centered care by ensuring smooth transitions among care providers.

Ultimately, this assessment instrument will support CMS' vision for a broader model of PAC in which payment, care-planning, quality assurance, and program monitoring are driven according to the individual beneficiary's needs. A more integrated approach to patient assessment will give providers in different settings more uniform information about a beneficiary's previous medical and health history to better plan and manage patient care.

However, it is important to recognize that assessment in and of itself is not the key to achieving this broader vision. Leaders of this effort need to avoid the pitfall of, "all assessed up and nowhere to go". In other words, avoid the scenario whereby the emphasis is disproportionately placed on assessment rather than initiating the appropriate actions to improve care. Assessment that does not directly address one of the three primary purposes described above could potentially adversely affect quality if clinicians' attention is distracted away from the essential task of preparing beneficiaries to participate in the care delivered in the next setting. Similarly, if the core information collected and transferred with the beneficiary is inaccurate, this approach could adversely affect patient safety by propagating medical errors. As with most successful efforts, this broader vision needs to articulate the desired outcome(s) first, and then work backwards to define how assessment can help to produce these outcomes.

Equally as important as defining what this assessment instrument would be is to concurrently articulate what this assessment instrument would <u>not</u> be. It would not replace existing setting specific assessment tools such as the MDS, OASIS, or IRF-PAI. It would not comprise a comprehensive data set for care planning. Rather, the assessment instrument would contain core data that could be incorporated into setting specific assessments and then augmented with additional information pertinent to a given setting.

In approaching the development of a uniform assessment instrument, a few lessons from the literature and a few caveats are worth considering. A recent Cochrane review concluded that there is insufficient evidence to suggest that hospital discharge planning reduces hospital readmission rates, hospital length of stay, hospital costs or that it improves health outcomes.[1] Similarly, beyond a few select conditions such as stroke and hip fracture, there is a paucity of research to assist clinicians in determining the optimal discharge destination setting for most beneficiaries. Clinicians in most care settings function in silos, operating independently from one another, and consequently have little insight into how care is delivered in other settings.[2] The different settings often have different staffing ratios, different orientations and approaches to assessment and management, and face different regulatory and reimbursement requirements. These different requirements also demand different data needs, and guidance will be needed for how the proposed uniform assessment instrument tool completed at hospital discharge would impact the MDS, OASIS, and IRF-PAI. For instance, the new instrument could either populate or be populated by these setting-specific assessment tools.

The timing for CMS to develop the uniform assessment instrument is ideal, as this effort will likely result in potential synergies with many existing national efforts. As will be discussed later in this chapter, new Joint Commission on Accreditation of Healthcare Organizations (JCAHO) initiatives will focus on continuity and coordination out of the hospital. The National Quality Forum has issued a call for performance measures aimed at coordinating care out of the hospital. The Institute for Healthcare Improvement Hospital is launching a transition out of the hospital initiative as part of its program aimed at Transforming Care At the Bedside. With support from the John A. Hartford Foundation, the Society for Hospital Medicine (SHM) has undertaken a consensus effort to create a model for Idealized Hospital Discharge Planning. The United Hospital Fund is planning a new program aimed at improving communication and collaboration between health care practitioners and family caregivers at the time of hospital discharge. Finally, the Institute of Medicine's December 2005 report "Performance Measurement: Accelerating Improvement", explicitly addresses transitional care and pay-for-performance.[3]

B. FINDINGS/DISCUSSION

1. Factors Influencing Beneficiaries' Post-Hospital Destination

Before discussing the potential application of a uniform PAC assessment instrument, it is worth considering the wide and varied factors that have the potential to influence a beneficiary's post-hospital destination. Many key factors may not routinely be captured during assessment. Ideally, the post-hospital destination decision should be a joint decision made between the responsible clinicians and the beneficiary, with input from family caregivers. The decision should be based primarily on the beneficiary's care needs at the time of transition and the beneficiary's preferences. The beneficiary should be informed as to what his or her financial obligation would be for the various PAC options under consideration.[4] To the extent that family caregivers are expected to assist in the execution of the care plan, their input should be explicitly obtained to determine whether they are willing and able to participate.

A host of factors that may not directly concern the beneficiary's immediate care needs often enter into the decision. Such factors include: service capabilities and staffing of the receiving

provider; complexity of patient care and resource needs; payor source(s); existing financial arrangements/contracts between providers; discharge planner's familiarity, professional and personal views; time of day, day of week; availability of able and willing caregiver support; patient/family preference; cultural preferences; patient cognition/capacity for carry over learning; transportation/geography/weather; and fear of litigation. Some of these factors are within the control of the clinicians handling the discharge, while others (e.g., daily per diem rates under the RUGS III system) are not.

Currently, there are no evidence-based criteria to inform hospital personnel in determining what the appropriate PAC setting is for a beneficiary with a known set of conditions and skilled care needs. Most of the studies conducted to date have focused on discreet conditions such as stroke and hip fracture. For example, one study by Kramer and colleagues found that since implementation of PPS, beneficiaries are increasingly transferred to IRF with the majority subsequently receiving care in a second Post Acute Care provider.[5] Previous work has demonstrated that beneficiaries discharged to inpatient rehabilitation facilities had improved functional outcomes compared with those discharged to SNF. However, this relationship was not observed for beneficiaries who suffered a hip fracture.[6] The researchers then used propensity scores to determine the factors that distinguished patients transferred to IRF versus SNF. Among numerous characteristics, the factors that distinguished IRF from SNF patients who suffered a stroke, the main factors were caregiver availability, cognition and social/recreational functioning. For patients who suffered a hip fracture, these same factors also predicted placement, as did physical function and walking ability. In developing guidelines for stroke care, the Agency for Health Care Policy and Research (now the Agency for Healthcare Research and Quality) was unable to reach consensus on the setting or service needs.[7] Thus the evidence needed to guide placement decisions is relatively limited and it is not clear that findings for one condition such as stroke can necessarily be generalized to other conditions.

2. Current Regulations That Might Impact a Discharge Planning Tool

Prior to the development of a uniform PAC assessment instrument, it may be instructive to consider how such a tool might be informed by and aligned with existing regulations. Medicare Conditions of Participation and HIPAA regulations are perhaps most relevant to the primary purposes articulated for this assessment instrument.

Within Medicare's statutory framework, Conditions of Participation explicitly include requirements concerning continuity of care and discharge planning for hospitals. JCAHO has deemed status from CMS to provide oversight for these Conditions of Participation. Under statute, this requires JCAHO to assess the care practices delineated in Table 4.1.

Table 4.1: Medicare Conditions of Participation Concerning Hospital Discharge Activities

Discharge planning: General Requirement
The hospital must have in effect a discharge planning process that applies to all patients. The policies and procedures for discharge planning must be specified in writing.

Identify Patients At-Risk for Adverse Health Consequences
The hospital must identify at an early stage of hospitalization all patients who are likely to suffer adverse health consequences upon discharge if there is no adequate discharge planning.

Discharge-Planning Evaluation
The hospital must provide a discharge-planning evaluation to the patients identified as at-risk and to other patients upon the patient's request or at the request of a physician.

Elements of the Discharge-Planning Evaluation
The discharge-planning evaluation must include an evaluation of the likelihood of a patient's needing post-hospital services and of the availability of the services.

Evaluating the Likelihood of Self-Care
The discharge-planning evaluation must include an evaluation of the likelihood of a patient's need for self-care or the possibility of patients being cared for in the environment from which they entered the hospital.

Timely Discharge Planning Required
The hospital personnel must complete the evaluation on a timely basis so that appropriate arrangements for post-hospital care are made before discharge and to avoid unnecessary delays in discharge.

Documentation of Discharge Planning and Patient Discussion
The hospital must include the discharge planning evaluation in the patient's medical record for use in establishing an appropriate discharge plan and must discuss the results of the evaluation with the patient or individual acting on his or her behalf.

Qualified Personnel for Discharge Plan Development
A registered nurse, social worker, or other appropriately qualified personnel must develop, or supervise the development of, a discharge plan if the discharge planning evaluation indicates a need for a discharge plan.

Physician Request for Discharge Plan
In the absence of a finding by the hospital that a patient needs a discharge plan, the patient's physician may request a discharge plan. In such a case, the hospital must develop a discharge plan for the patient.

Hospital to Arrange Services
The hospital must arrange for the initial implementation of the patient's discharge plan.

Reassessing the Discharge Plan
The hospital must reassess the patient's discharge plan if there are factors that may affect continuing care needs or the appropriateness of the discharge plan.

Pre-discharge Counseling
As needed, the patient and family members or interested persons must be counseled to prepare them for post-hospital care.

Transfer and Referral
The hospital must transfer or refer patients, along with necessary medical information, to appropriate facilities, agencies, or outpatient services, as needed, for follow-up or ancillary care.

JCAHO has been criticized for its lack of vigorous oversight and monitoring of the Medicare hospital Conditions of Participation in general[8] and with respect to the discharge planning process in particular.[3;4] In 2002, more than 90% of all hospitals nationwide received the highest score of 5/5 (i.e., "substantial compliance") for discharge planning accreditation items.[3] These findings are in sharp contrast to the growing evidence base that demonstrates there are serious quality problems in transitional care.[9] In its defense, JCAHO has explained that it believed it does not have "deemed status" with respect to compliance with discharge planning standards (i.e., that these particular Conditions of Participation had been "carved out"). However, subsequent discussions between CMS and JCAHO have clarified that in fact JCAHO does have deemed status for discharge planning standards.

Among those Conditions of Participation related to discharge planning identified in Table 4.1, the standard for a Discharge Planning Evaluation is particularly relevant to the development of a uniform PAC assessment instrument. The discharge-planning evaluation is different from the discharge plan. The evaluation is an assessment that looks at the patient's physical and mental condition, the likely post-hospital living situation, and the patient's ability to engage in such daily living activities as eating, dressing, bathing, and ambulating. The plan, including the type of setting to which the patient is to be discharged, focuses on the medical and social support needs of the patient in that setting.

CMS' Interpretive Guidelines provide that the needs assessment can be formal or informal. The hospital may develop an evaluation tool or protocol. Generally, the assessment should include an evaluation of factors that affect an individual's needs for care after discharge from the acute care setting, such as an assessment of bio-psychosocial needs, the patient's and caregiver's understanding of discharge needs, and identification of post-hospital care resources. At the present time, nonetheless, there is no nationally accepted standard for this evaluation. The purpose of a discharge planning evaluation is to determine continuing care needs after the patient leaves the hospital setting. It is not intended to be a care-planning document.

Returning to JCAHO, in January 2004, JCAHO implemented a new approach to the survey process, Tracer Methodology.[10] This new approach also has potential implications for the development of a uniform PAC assessment instrument. Tracer Methodology includes the following elements: (a) following the course of care and services provided to a particular patient; (b) assessing relationships among disciplines and important functions; (c) evaluating the performance of relevant processes related to patient care; and (d) identifying potential vulnerabilities in care processes. It is now part of the typical 3-day onsite hospital survey process, and in most instances, a typical team of three surveyors is expected to complete approximately 11 tracers. The Tracer Methodology has not yet been extended beyond the hospital setting but it has potentially important implications for discharge planning and transitions. In particular, this approach can follow a particular patient, assessing how the patient fares along a continuum of care. It can potentially assess how well the hospital staff has ascertained post-hospital needs of a particular patient, the planning for discharge that has occurred, and, through patient interviews, can assess the patient's understanding about the PAC aspects of his or her care. At the time of this writing, JCAHO plans to extend the Tracer Methodology to include the discharge process from hospital to home. JCAHO surveyors will

contact beneficiaries several days after discharge to assess whether hospital staff made adequate arrangements and appropriately prepared beneficiaries for discharge.

Another particularly relevant regulatory concern that may potentially influence the development of a uniform PAC assessment instrument is the Health Insurance Portability and Accountability Act (HIPAA). Exchange of patients' health information for purposes of treatment, payment, and health care operations is allowed under HIPAA. This would include not only the physicians and nursing staff but also case managers, home and community services personnel, and the communication of referral information to multiple facilities just prior to the decision determining appropriate placement. Thus although HIPAA regulations are relatively clear, misinterpretation is common. Clinicians' may be apprehensive to share information for fear of violating HIPAA laws and facing accompanying penalties. The HIPAA regulations are not the primary barrier per se but rather it is the perceptions, the myths, and the false assumption about the regulations. If CMS were to implement a uniform PAC assessment, it would be important to take this opportunity to provide accompanying language clarifying CMS' interpretation of HIPAA laws for potential end-adopters.

3. Existing Hospital Discharge Assessment Tools

Prior attempts to create a uniform PAC assessment instrument can provide important lessons and insights to the current discussion. In some cases, the content of the instrument under development provides valuable information into the domains of the new instrument proposed, which are discussed in the following section. However, in other cases, it is the set of actions or processes that did or did not lead to successful implementation that are particularly instructive. Four instruments will be discussed: the Uniform Needs Assessment Instrument (UNAI), the Veteran's Association Geriatrics and Extended Care (GEC) Referral, a predictive tool developed by Holland and colleagues, and a prediction tool developed by Fairchild and colleagues.

The creation of the UNAI was mandated under OBRA 87. Not unlike the current effort discussed herein, the primary purpose of the UNAI was to determine individual's needs for continuing care and not to replace comprehensive geriatric assessment.[11] The UNAI was intended to improve continuity of care between acute, post-acute, and long-term care. It was explicitly designed to help discharge planners make more informed decisions and also transmit essential information to PAC providers to initiate appropriate and timely care. An extensive effort was made to create the UNAI and the domains that were selected are quite similar to those proposed in the next section for the new instrument. Research Triangle Institute and Meyers Research Institute received the contract to conduct field-testing.[12;13] Thus there are many parallels between this past effort and the current effort under discussion. According to CMS (HCFA) staff working directly on this instrument in the 1990's, the reasons why the UNAI effort did not go forward were two-fold: (1) there was no Congressional mandate for implementation; and (2) the UNAI was not tied to payment. To the extent possible, this current effort will be strengthened by directly addressing both of these critical issues that were significant barriers to the success of a similar prior effort.

The VA Geriatrics and Extended Care (GEC) Referral Discharge Planning Tool also provides considerable overlap with the current effort towards the creation of a uniform PAC assessment instrument. The VA GEC referral form contains items pertinent to determining the level of care

required by the patient. It was designed to serve as a single document that could be shared with multiple PAC providers, thereby facilitating a way that a common set of information would be available to all programs that might accept the patient. The GEC referral form is not intended to be a full assessment; rather, it is a screening tool. The tool was not designed for outcome assessment, so some of the measures such as ADL and IADL are gross scales that do not map onto single metrics and would be unlikely to be sensitive to change. Although it shares many of the domains identified as important for the Medicare assessment instrument, the form deliberately does not include information that can be automatically retrieved from VISTA, the VA's electronic health information system. Such information includes a problem list, medication list, orders, notes, consults, discharge summary, labs, reports, and vital signs. A 2004 study of interoperable electronic health records for PAC and long-term care settings by Kramer et al. found that a major limitation of the VA's VISTA system is that it does not extend to providers outside of the VA system.[14] This is particularly problematic in relation to medication data, which were at times found to be incomplete, even when the non-VA settings used electronic systems. Because transfer of medication information is especially important in facilitating appropriate PAC placement and transitions between PAC settings (see discussion below), the inability to transfer this information across PAC settings is of concern. Current testing of the VA GEC referral form is underway in the Pacific Northwest and the VA is planning a national rollout in early 2006.

Holland and colleagues from the Mayo Clinic have developed and validated a screen for specialized discharge planning services.[15] These services were operationally defined as multifaceted and complex, requiring coordination between hospital and community-based care providers. Of 24 different variables examined, only four (age, living alone, disability, self-rated walking limitation) predicted use of specialized discharge planning services with sensitivity of 75% and specificity of 78%. The developers are currently exploring how to obtain the variables directly from the electronic health information system.

Fairchild and colleagues created a prediction rule for the use of post-discharge medical services.[16] These services were defined as use of visiting nurses or physical therapists, medical equipment, or placement in a rehabilitative or long-term care facility. Similar to the findings of Holland and colleagues, the prediction tool uses age and SF-36 physical function and social function scores to stratify patients with regard to their likelihood of needing discharge planning to arrange for post-discharge medical services. While the work of Holland and Fairchild may not directly pertain to the current effort discussed in this report, the variables identified may have application to the next section on identifying the appropriate domains.

4. Candidate Domains for Inclusion in an Assessment Tool

Before discussing potential candidate domains to be included in a uniform PAC assessment instrument, it is important to re-emphasize several points made at the beginning of this chapter. First, the three primary purposes for this tool (facilitate appropriate post-hospital placement, provision of core patient information to promote of safe and efficient care transitions between care settings, and assessment of outcomes across settings and over time) must drive the selection of the domains. In other words, the developers need to begin with the desired outcomes and work backwards. Second, this tool will not be comprehensive in scope or content. Doing so

would likely guarantee a failed implementation and this new tool would end up on the shelf. Individual post-acute and long-term care settings will need to supplement the information contained in this tool with their own setting-specific information. This includes information used to complete the MDS, OASIS, and IRF-PAI. It is conceivable that items from these instruments could be used to populate this new instrument and vice versa. Third, the construction of this new instrument needs to be beneficiary-centered, accounting for beneficiary preference(s) and beneficiary financial obligation, and it needs to acknowledge the central role assumed by family caregivers.

Table 4.2 summarizes candidate information that hospital staff routinely rely upon to facilitate appropriate post-hospital placement. Earlier in this chapter, a number of factors beyond the purview of the hospital staff were articulated. The items included in this table are largely within the purview of hospital staff. This is not to say that each item is already gathered or used in a systematic way. For example, the first item, goals of care, is often not explicitly stated and often lacks the essential input of beneficiaries and their family caregivers. Further, articulating a beneficiary's skilled care needs requires a broader understanding than simply knowing what transpired during the hospital stay. It necessitates taking the time to learn what services the patient needed and among those, which were received prior to hospital admission. It also includes conferring with family caregivers to determine if they are able and willing to sustain their current level of involvement, before giving them additional responsibilities.

Table 4.2: Assessment to Determine Discharge Needs/Setting

- Goals of care—rehabilitation, skilled nursing, monitoring, respite, palliation, reduce hospital visits/adherence, stabilize behavior
- Skilled nursing or rehab care needs (RN, PT, OT, ST) (Oxygen, suctioning, tracheostomy care, ostomy care, tube feeding, parental feeding, IV infusions, medications by injection, urinary catheter care, dialysis, wound care, pressure ulcer care)
- Patient's current residence and with whom
- Meets Medicare criteria for homebound (y/n)
- Current services receiving in home (pre-hospital)
- Current DME receiving (pre-hospital or in hospital)
- Capacity to perform self-care (includes cognition/carry over learning)
- IADL supportive needs/ADL supportive needs (pre-morbid and current)
- Pre-morbid and current cognitive functional status
- Pre-morbid and current physical functional status
- Family caregiver willing and able to help with the execution of the care plan as outlined? What is caregiver relationship to patient, what support currently provided (advice/emotional support, ADL help, IADL help, medication administration, and transportation)? Does caregiver live with patient?
- Ethnic or cultural considerations
- Infectious precautions
- Active problem list (including behavioral problems)
- Medication list (as it pertains to per diem)
- Prognosis/estimated duration of services
- Insurance/payor

As discussed earlier, the factors that determine discharge placement are not well understood and nationally wide variation has been observed. Prediction tools are being developed but much work remains. Nevertheless, in the absence of an extensive research base, these decisions are being made thousands of time per day and a uniform PAC assessment instrument could help to bring greater standardization to an otherwise heterogeneous process. The domains listed in Table 4.2 could serve to drive beneficiary placement to appropriate PAC settings. (Please note that the functional domains listed in this chapter are defined more specifically in Chapter 6.)

For example, a beneficiary who suffered a stroke whose goal is primary rehabilitation and who has the functional reserve to withstand intense rehabilitation would likely be best served in an IRF. In contrast, a beneficiary with advanced congestive heart failure whose goals are primarily palliative with transition to end-of-life care might be discharged to a skilled nursing facility for initiation of hospice care. As another example, a beneficiary whose skilled care needs primarily included intravenous antibiotics and wound care for an infected lower extremity with secondary rehabilitative needs might be preferentially discharged to a skilled nursing facility. Finally, a patient with intact cognitive function who prior to hospitalization was living in her own home and managing her self-care without difficulty may return home with skilled home care services and daily visits from a concerned family caregiver.

Table 4.3 includes a suggested set of domains that could be used to ensure provision of core patient information to promote safe and efficient care transitions between care settings. The goal is to provide the receiving team of clinicians with a core set of information needed to provide high quality and safe care of the beneficiary. To reiterate, the domains included in Table 4.3 are but a subset of all possible items that clinicians in acute, post-acute and long-term care settings may use in constructing a comprehensive care plan for an individual beneficiary. The domains in this table represent a starting point and undoubtedly, modifications will be made. This table was directly informed by the domains included in the UNAI, the VA GEC, a Colorado-based collaborative project between hospitals, PAC providers, payors, and the QIO, and from a best practices report on care transitions based on the work of some of the nation's exemplary health delivery systems.[17]

The information included in Table 4.3 is designed to help the receiving care team quickly develop an understanding of an individual beneficiary from the standpoint of functional status (pre-morbid as well present), the immediate care needs that require attention, an accurate and up-to-date medication list, self care ability and family support, and insurance coverage. All too often an older adult exhibits some confusion related to an acute illness or a medication side effect and is incorrectly labeled as having Alzheimer's Disease because key information such as this is unavailable.

Once the potential domains for a uniform PAC assessment instrument are selected, an important next step will be to identify the most accurate and efficient sources of data collection to populate these fields. Administrative data is often the most efficient to use and can reduce provider burden but suffers from the fact that it does not include the majority of information in Tables 4.2 and 4.3. Clinician assessment has been the most commonly used approach for hospitals using their own assessment instruments. Accountability for the completion and accuracy of this information has been a gray area. Finally, patient self-report would ultimately provide the most

useful information for promoting beneficiary-centered care and would ensure that patient preferences and family caregiver input is obtained. However, due to transient or progressive cognitive impairment, not all beneficiaries are in a position to participate in such an assessment.

Table 4.3: Core Information Set Needed for Transfer

- Goals of care—rehab, skilled nursing, monitoring, respite, palliation, reduce hospital visits/adherence, stabilize behavior
- Active problem list (including depression/anxiety)
- Medication list—reconciled—including indication for each medication and once reconciled, indicate new medication, same medication different instructions, same medication same instructions, medications suspended (with guide when to re-evaluate), medications stopped
- Allergies/intolerances
- Resuscitation status/advance directive/DPAHC
- Discharge instructions/outstanding diagnostic tests/scheduled follow-up appointments
- Pre-morbid and current cognitive functional status
- Pre-morbid and current physical functional status
- Pain status
- Skin integrity
- Sensory deficits (vision, hearing, neuropathy)
- Dietary needs
- Continence
- Fall risk
- Current services receiving in home (pre-hospital)
- Current DME receiving (pre-hospital or in hospital)
- Capacity to perform self-care (includes cognition/carry over learning)
- IADL supportive needs/ADL supportive needs
- Family caregiver willing and able to help with the execution of the care plan as outlined? What is caregiver relationship to patient, what support currently provided (advice/emotional support, ADL help, IADL help, medication administration, and transportation)? Does caregiver live with patient?
- Ethnic or cultural considerations/language
- Equipment/assistive devices
- Immunizations (Pneumonia, Influenza, Tetanus) and most recent PPD testing result
- Self-rated health status
- Insurance/payor

5. Communicating the Assessment

With the domains selected and the source of information identified, the next question becomes how will the assessment be communicated from the hospital to PAC provider? The inadequacy of the current non-standardized approach poses problems for both quality and safety.[17;18] Within our existing technology, the current options include communicating the information electronically (through interoperable electronic health information systems or e-mail), facsimile, U.S. Mail, or sending a paper copy of the assessment with the beneficiary.

Similar to other national quality improvement and performance measurement initiatives, implementation of a uniform PAC assessment instrument could also serve to promote more widespread use of health information technology. To date, most electronic health information systems are based in hospital or ambulatory settings with little to no extension into the post-acute and long-term care arenas.[19] Ensuring interoperability across disparate information systems remains a significant challenge but could be overcome with the proper motivation and financial incentives that would stimulate providers' interest in making this investment. However, before existing software can support auto-population of comparable fields to promote cross-site data transfer, standards will need to be developed and approved for many of the items included in Tables 4.2 and 4.3.

Beyond the transmission of information in a timely accurate manner, promoting the use of health information technology potentially opens the door for the use of decision support software that may be able to identify high-risk beneficiaries and facilitate their placement to PAC settings based on the current state of evidence. Health information technology could facilitate the assessment of outcomes across settings and time, thereby creating the opportunity for performance measurement and pay-for-performance initiatives.

6. How a Hospital Discharge Assessment Tool Could Support Pay-for-Performance

The Institute of Medicine's recently released report on Performance Measurement encourages CMS to address the current measurement gaps, including the need for comprehensive measurement, longitudinal measurement, patient-, population-, and system-level measurement and shared accountability. The IOM report specifically identifies measures of continuity and transitions as priority areas. Thus, in many respects, the IOM report has set the stage for how a uniform PAC assessment instrument might support a value-based purchasing or pay-for-performance initiative. The National Quality Forum has issued calls for measures of care coordination out of the hospital and in ambulatory care as well.

A gradual or phased-in pay-for-performance approach may be preferred over a more ambitious full-scale implementation. For example, providers might first be given a differential payment for simply completing the PAC assessment instrument. Next, the bar would be raised such that the differential payment would only be issued for the accurate completion of the assessment. This is particularly important for transfers between acute and PAC providers. For example, there are often problems with accountability and accompanying "finger-pointing" about skin integrity whereby each provider blames the other for a new pressure ulcer. Further, as with any assessment, there is a risk that the clinician may indicate that an assessment has been completed without actually completing the assessment. This is particularly important for activities that assess a beneficiary's capacity for self-care. Feedback loops that include beneficiary self-report will be needed. Next, the added payment would only be made if the information from the assessment were transmitted in a timely manner. Finally, payment would be made for the achievement of outcomes that are based upon the uniform PAC assessment instrument. Using this approach, CMS leaders and Congress will have a better understanding of the value of services paid for by the CMS budget.

7. Facilitating the Implementation of a Hospital Discharge Assessment Tool

The successful implementation of a uniform PAC assessment instrument across acute, post-acute, and long-term care providers will ultimately hinge on obtaining the support and "buy-in" from provider groups and professional societies. Clinicians working in these settings are overextended and mandating a new assessment that is perceived to not directly benefit patient care will be met with strong resistance. It is essential that this new assessment is not viewed as simply a check off list but rather be seen as an interactive process that promotes healing relationships. It is also critical that this new assessment does not detract from clinicians' face-to-face time with beneficiaries, as this is already inadequate for supporting post-hospital self-care activities.

Thus it is worth considering the potential advantages a uniform PAC assessment instrument may offer. From the standpoint of the hospital, implementing a standardized assessment tool that facilitates appropriate placement could potentially expedite discharge planning, resulting in an overall reduction in hospital length of stay. With increasing attention and rigor planned for future JCAHO surveys, an assessment tool may enhance a hospital's accreditation. Many hospitals have embraced patient safety initiatives and the timely and accurate communication of a core set of clinical information could be encouraged as a patient safety goal. Additional advantages might include less redundancy in assessment, presuming that for beneficiaries re-admitted to the same facility, the assessment could be updated rather than began anew. Hospitals may also see the value in having to only complete a single assessment for all potential PAC transfers, rather than having to fill out a new form for each potential provider group. Finally, should the IHI transition out of the hospital program prove successful, CMS could work with the QIOs to promote widespread adoption of this model.

D. REFERENCES

1. Parkes J, Shepperd S. Discharge planning from hospital to home (Cochrane Review). Cochrane Database of Systematic Reviews. 3. 2003.

2. Coleman EA, Berenson RA. Lost in transition: Challenges and opportunities for improving the quality of transitional care. Ann Intern Med 2004; 141(7):533-536.

3. Coleman EA. Background paper on transitional care performance measurement. Institute of Medicine Performance Measurement - Accelerating Improvement. 2005. Washington, D.C., The National Academies Press.

4. Chiplin JrAJ. Breathing life into discharge planning. The Elder Law Journal 2005; 13(1):1-83.

5. Kramer A, Goodrich G, Holthaus D, Epstein A. Chapter 4: Patterns of post-acute care following implementation of PPS (Final Draft Report). Study of Stroke Post-Acute Care and Outcomes. 2006. Division of Health Care Policy and Research, University of Colorado at Denver and Health Sciences Center.

6. Kramer A, Steiner J, Schlenker R, Eilertsen T, Hrincevich C, Tropea D et al. Outcomes and costs after hip fracture and stroke. JAMA 1997; 277(5):396-404.

7. Gresham GE, Duncan PW, Stason WB, et al. Post-stroke rehabilitation. Clinical Practice Guideline, No. 16. 1995. Rockville, MD, U.S. Department of Health and Human Services. Public Health Service, Agency for Health Care Policy and Research. AHCPR Publication No. 95-0662.

8. Inspector General. Medicare hospital discharge planning. Office of Inspector General, Department of Health and Human Services 1997; http://www.oig.hhs.gov/reports.html(OEI-02-94-00320).

9. Coleman EA, Berenson RA. Lost in transition: Challenges and opportunities for improving the quality of transitional care. Ann Intern Med 2004; 141(7):533-536.

10. Joint Commission on Accreditation of Healthcare Organizations. Shared visions - New pathways Q&A. New pathway: Tracer methodology. http://www.jcaho.org/accredited+organizations/svnp/svnp+qa_tracer+methodology.htm Accessed 1/1/05. 2004.

11. Health Care Financing Administration (HCFA). Advisory panel to develop uniform needs assessment instrument(s). U.S.Department of Health and Human Services Report to Congress: Report of the Secretary's Advisory Panel on the Development of Uniform Needs Assessment Instrument(s) , 1-253. 1992. Washington, D.C., U.S. Government Printing Office.

12. Westra BL, Holland DE, Aufenthie J, Cullen L, Finley J, Griebenow L et al. Testing the Uniform Needs Assessment Instrument for hospital discharge planning with older adults. Journal of Gerontological Nursing 1998; 24(5):42-46.

13. Research Triangle Institute and Myers Research Institute Project 5760. Health Care Financing Administration (HCFA) Field Testing of the UNAI (Uniform Needs Assessment Instrument) Small Scale Trial 2 - Data Collection Manual. 1-152. 1999.

14. Kramer A, Bennett R, Fish R, Lin CT, Floersh N, Conway K et al. Case studies of electronic health records in post-acute and long-term care. 4-18-2004. U.S. Department of Health and Human Services, Assistant Secretary for Planning and Evaluation, Office of Disability, Aging and Long-Term Care Policy.

15. Holland DE, Harris MR, Leibson CL, Pankratz VS, Krichbaum K. Development and validation of a screen for specialized discharge planning services. Nursing Research. In press.

16. Fairchild D, Hickey M, Cook E, McCarthy R, Rossi L, Timmons S et al. A prediction rule for the use of postdischarge medical services. J Gen Intern Med 1998; 13(2):98-105.

17. Coleman EA, Fox PD, on behalf of the HMO Care Management Workgroup. One Patient, Many Places: Managing Patient Transitions of Care. http://www aahp org/Content/NavigationMenu/Inside_AAHP/Care_Management1/Care_Management htm 2004.

18. Coleman EA, Berenson RA. Lost in transition: Challenges and opportunities for improving the quality of transitional care. Ann Intern Med 2004; 141(7):533-536.

19. Kramer AM, Bennett R, Fish R, Lin CT, Floersch N, Conway K et al. Case Studies of Electronic Health Records in Post-Acute and Long-Term Care - Final Report. Office of Disability, Aging, and Long-Term Care Policy. Department of Health and Human Services[Contract Number:233-02-0070], 1-141. 8-18-2004.

Chapter 5
Longitudinal Outcome Monitoring across PAC Settings

A. INTRODUCTION AND BACKGROUND

As discussed in Chapter 1, the Centers for Medicare and Medicaid Services (CMS) has acknowledged that steps need to be taken toward developing a PAC system in the United States that provides payment and assures quality for an overall episode of PAC, rather than for each individual component within the continuum of care. As an essential step toward accomplishing this policy objective, assessment methods are needed to collect and compare relevant outcome and quality of care data across various sites where PAC is provided.

Unfortunately, since each PPS system within PAC was developed independently for each PAC setting, patient-centered assessment systems used as the informational platforms for these new payment systems are distinct and setting specific. Consequently, a persistent barrier to fulfilling this policy mandate for PAC has been the inability to achieve a standardized, patient-centered outcome assessment approach that can provide CMS and other interested stakeholders with appropriate information on outcomes and quality of care that can be applied over time, and across different settings where PAC services are provided.[4]

The purpose of this chapter is to introduce the reader to contemporary methods for developing and evaluating outcomes measures which, we believe, provide the CMS with a methodology that can be implemented to track outcomes and quality of care provided across entire episodes of PAC. In the first section of this chapter we will summarize the major limitations in traditional methodology for assessing health outcomes. We will then briefly introduce contemporary measurement concepts of Item Response Theory (IRT) and Computer Adaptive Testing (CAT) methods as alternatives to the traditional approaches used to assess and monitor PAC outcomes. Finally, we will explain how contemporary techniques for measuring outcomes can be used as a strategy to achieve episode-wide assessment in PAC and discuss some recommendations to CMS for implementing episode-wide outcomes assessment and monitoring in PAC.

B. SCOPE OF VARIABLES

To recap, there are three primary purposes for uniform PAC patient outcomes assessment that are the focus of this report: 1) assessment at the point of hospital discharge for the purpose of achieving appropriate PAC placement; 2) facilitating care transitions across PAC settings; and 3) outcome assessment for longitudinal monitoring of quality of care.

The focus of Chapter 5 is on methods for achieving longitudinal outcome assessment for the purpose of monitoring and improving the quality of care provided across an episode of PAC. We will not discuss assessment for the purposes of achieving appropriate PAC placement or for facilitating care transitions within an episode of PAC; however, we do believe that the contemporary measurement techniques described in this chapter may also have important applicability for assessments aimed at hospital discharge placement and to facilitate care transitions across PAC settings.

In this chapter, we will use the example of functional status for the purpose of illustrating the application of contemporary measurement methods to PAC. In putting forth recommendations on the need and potential for standardizing patient-centered data collection for the purposes of quality monitoring and quality assurance purposes, the National Committee on Vital and Health Statistics(NCVHS)[5] was unanimous in recommending information on a *patient's functional status* for inclusion in such quality monitoring systems. Consistent with the NCVHS recommendations, in this chapter we use the term functional status very broadly to include an individual's ability to carry out activities of daily living (ADLs and IADLs) as well as their participation in advanced functional activities.

C. SHORTCOMINGS OF TRADITIONAL MEASUREMENT METHODS

Outcome systems in PAC traditionally take the form of analyses at a single-item or multiple-item (instrument) level. A single-item approach employs a functional status question that stands alone and can estimate a patient's functional status at a point in time or data on functional change. Single items have been identified, for example, on the OASIS as target outcomes for a given clinical sample.[6] Although a single item outcome approach has the advantage of simplicity of use and direct interpretability,[7] it has important limitations. Single items can be much less reliable than an instrument comprising multiple items, and since outcomes such as functional status outcomes are complex, they may not lend themselves to valid measurement and tracking using a single item.[8] With a multiple-item instrument, each item is designed to sample content important to the underlying outcome, leading to improved reliability, validity, precision, and responsiveness in comparison to a single item.

There are numerous well-respected, multiple-item functional outcome instruments in widespread use within PAC.[9-13] Some have developed aggregate scoring systems and sub-domain scores by summing raw scores (such as the FIM), and others are primarily analyzed at the single-item level (such as MDS).[14] A barrier toward developing meaningful summary scores for certain multiple-item instruments used in PAC (such as the OASIS) is the varying response scales across OASIS items, which range from 4 to 6 categories. Using conventional methods, certain items may be weighted more based on a greater number of response categories.[15]

Most existing functional status instruments that generate and employ summary scores have been created under the umbrella of Classical Test Theory (CTT), a set of assumptions and psychometric procedures that have been widely used to develop tests for much of the 20th century.[16] The defining signature of functional status instruments developed using CTT is the use of a fixed set of questions (or items) presented to a respondent, regardless of the appropriateness of any specific item for a given PAC patient. [In this discussion we will use the term respondent when referring to the person who is completing the functional status assessment. The respondent can be a patient receiving care, one or more clinicians providing care, or a family member or significant other serving as a proxy respondent for a patient]. Items in most traditional instruments are selected or written to represent a moderate range of functions at a moderate level of difficulty or limitation. By selecting items that are fairly homogeneous, acceptable levels of reliability and validity with a limited number of items can best be achieved. The major limitation of CTT for application in PAC is the reliance on a fixed set of items (test-

dependency) that are expected to measure functional status across a wide variety of patients and settings.[17]

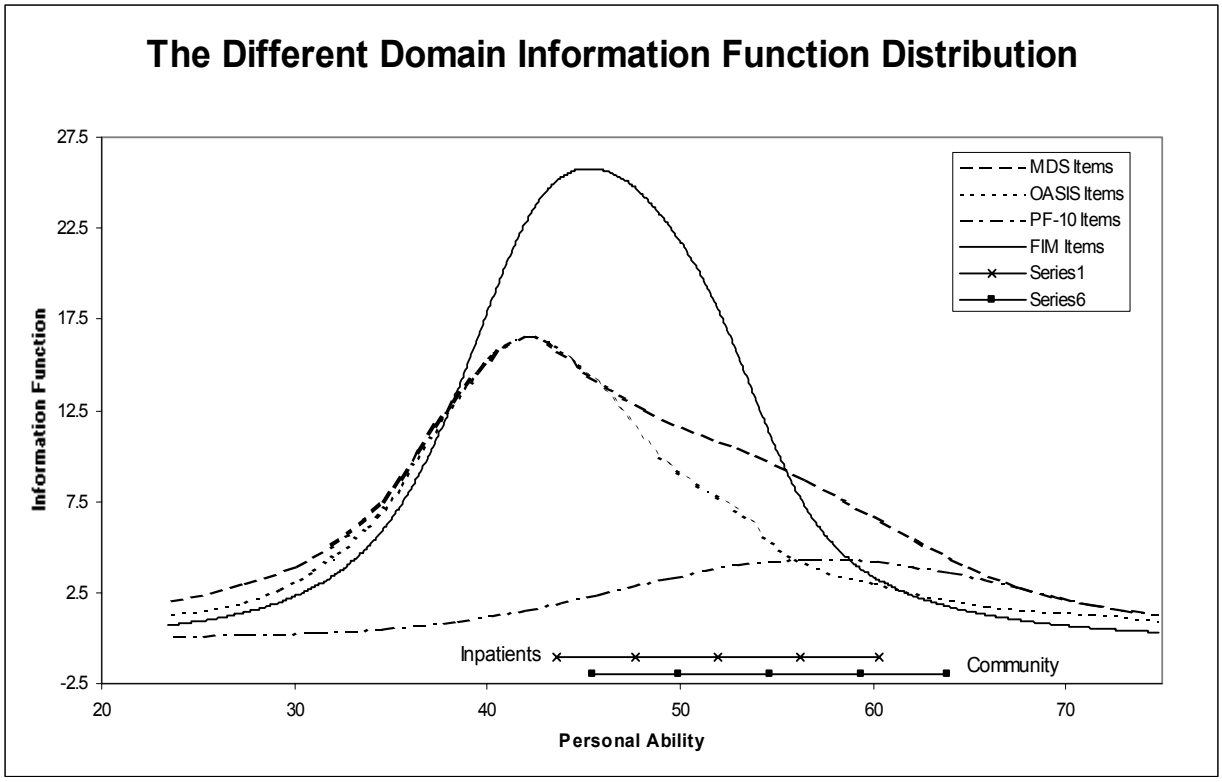
As was stated earlier, separate instruments have been developed for each PAC setting. Data incompatibility across PAC instruments renders the ability to track relevant outcomes across different care settings almost impossible to accomplish. For example, a recent attempt to rescale items from the MDS-PAC to FIM-like items in a desire to explore whether the MDS-PAC could be substituted for the FIM was not sufficiently accurate to ensure payment equity.[18] Because each instrument used a different set of items, rating scales, and different administration and scoring rules, payment cell classification using FIM data agreed with that using MDS-PAC data only 56% percent of the time.[19]

Instead of developing translational cross-walks across different functional status instruments, some have argued for the development of a core set of functional items that could be used to monitor functional outcomes across PAC settings.[4] To be useful, this core set would need to contain items that cover a wide range of the functional status continuum with good measurement precision at all levels.[20-22]

In a recent study, we assessed the test precision using test information function across four key setting-specific functional status instruments used within PAC.[23] This was done in an attempt to identify whether one traditional functional status instrument could be used for the expressed purpose of monitoring the quality of PAC over time, and across different settings where PAC is provided. In a sample of PAC patients, we collected functional status items from standardized instruments including, the 18-item Functional Independence Measure (FIM^{TM})[12] for persons in inpatient rehabilitation settings, 19 Minimum Data Set (MDS)[24] items (physical functioning and selected cognitive items) for persons in skilled nursing home settings, 19 ADL/IADL items from the Standardized Outcome and Assessment Information Set for Home Health Care (OASIS)[10] for persons receiving home care, and the 10 physical functioning items of the SF-36 for individuals receiving outpatient services.[25]

In Figure 5.1, the measurement precision of each instrument is depicted by its unique test information function curve. These curves are superimposed with the average score on the functional continuum and ±2 sd for the inpatients and community samples. Note the location of the peak amount of precision for each instrument in relationship to each sample. Although the OASIS items contain a broad range of content, they provide a high degree of measurement precision at only the very low end of functional status dimension for both the inpatient and community samples. The precision of MDS items is also greatest at the lower functional status dimension levels, although the MDS items have a greater span of functioning in which they provide some levels of precision. The FIM items peak at the low to moderate end of the inpatient sample. In contrast, the information function of the PF-10 items peaks at the high end of the community functioning sample with very poor precision for the inpatient sample.

Figure 5.1: Test Precision of FIM, MDS, OASIS, and PF-10



The results of these analyses illustrate the challenge of finding or constructing one instrument with sufficient precision across a broad range of functional status. In our judgment, none of the instruments examined in this study achieved an acceptable level of precision across the broad continuum of function for longitudinal monitoring of function in PAC. Because setting-specific functional status instruments are geared to the functional level of patients in a particular PAC setting, those patients at either end of the functional continuum covered by a particular instrument tend to be poorly defined. The FIM, for example, yields serious ceiling effects when applied to home care patients, while the PF-10 possesses serious floor effects when applied with more disabled populations. Ceiling and floor effects are serious concerns when functional status instruments are applied for quality of care monitoring. It is impossible to distinguish among patients at the ceiling or floor, even though they may be functionally different in important ways. For quality of care monitoring, ceiling effects produce serious type II errors in testing whether or not outcomes have improved in PAC patient populations, while it is equally difficult to determine if there have been important declines in function for persons at the floor.

A final comment on practicality should be mentioned. Using a CTT approach, it is impractical for any one instrument to include the number of items necessary to yield precise measures across the wide range individuals' function as they progress throughout an episode of PAC.[22] Patients and/or clinicians are often frustrated by being asked to respond to items that seem redundant, of low relevance, or take considerable time to complete.[26] The resulting length and complexity of such approaches may be problematic and raise concerns over respondent burden, administration

costs, and acceptability to users. The "ideal" measure, possessing enough questions to cover the full content range with a high degree of precision at all relevant levels, however, is impractical when using traditional measurement technology. Although the application of traditional functional status instruments have led to important insights into PAC practice, we believe the deficiencies inherent in CTT methodology have impeded progress toward achieving a successful uniform methodology for monitoring, managing, and improving the outcomes of services provided to PAC patients across the entire episode of care.[23;27]

D. A PARADIGM SHIFT: CONTEMPORARY MEASUREMENT METHODS

Precise measurement of functional status at all levels is a critical feature of any measurement method proposed for monitoring outcomes for the purpose of quality of care determination across the entire continuum of PAC. The assessment goal at each point in time is to match the level of item difficulty to the ability level of the patient at any particular point in the PAC episode, and thus provide a set of items in an assessment that closely corresponds to the current functional level of the patient.[17] There is a way to achieve feasible, practical, and precise measurement of functional status through the conjoint use of two Contemporary Measurement Theory methods, a test administration method known as ***Computer Adaptive Testing (CATs)*** coupled with a measurement theory known as ***Item Response Theory (IRT)***.[17;28-31]

1. Computer Adaptive Testing

CAT selects test items tailored to an individual patient, shortens or lengthens the test to achieve the desired precision of measurement, scores every patient on the same underlying outcome continuum so that results can be compared across the continuum of that outcome, and displays the results instantly for immediate interpretation and use. CAT administration requires computer technology to administer the instrument and is adaptive in the sense that each item administered is tailored to the unique outcome level of a patient. Each respondent is administered a different version of the instrument because individual items are selected based on the respondent's previous answers.

The basic notion of a CAT test is to mimic what an experienced clinician does. A clinician learns most when he/she directs questions at the patient's approximate level of proficiency. Administering outcome items that represent tasks that are either too easy or too hard for the patient provides little information. In contrast to traditional, fixed form functional tests that ask the same questions of everyone regardless of how the respondent answers, CAT instruments, like a skilled clinician, tailor their assessment by asking only the most informative questions based on a person's responses to previous questions.

A CAT is programmed to first present an item from the mid-range of a defined list of functional items, and then directs subsequent functional items to the level based on the respondent's previous responses, without asking unnecessary questions. The selection of an item in the mid-range is arbitrary and the CAT can be set to select an initial item based on other information entered about the patient such as age, diagnosis, or severity of their condition. By having comprehensive item banks available in each functional outcome domain of interest, the selection of additional items after the initial one is based on responses to the previous items. This allows

for fewer items to be administered while gaining precise information regarding an individual's placement along an outcome continuum.[32]

An important advantage of CAT is the ability to update the item bank and the CAT programs easily as new items become available. Since item banks are built using an IRT model, items can be updated, changed, or removed without disputing the overall functional scale and scoring metric. For example, if new items were written that have potential advantages over existing items, or cover content that is not currently in the item bank, those items could be imbedded in a CAT program. These new items would not be part of the current scoring system, but rather would be "test items." These new items could then be examined in an IRT analysis to determine if they fit the functional scale, have content advantages over current items, and have locations on a functional scale that fill in a content gap, and then subsequently be incorporated into the next revision of the CAT program.

Although relatively new in the health field, CAT methodology is already commonplace in knowledge-based testing situations. For example, the Graduate Record Examination, the Scholastic Aptitude Test, and the Test of English as a Foreign Language, as well as numerous professional credentialing and licensure examinations use CAT methodology. CAT methods reduce instrument administration times dramatically while allowing test precision to be tailored to the specific needs of a particular testing situation or context. Large educational testing organizations that provide services for college entrance are constantly testing new items in existing exams so that new items that better reflect the construct of interest are being developed for future assessments. If introduced into PAC assessment, we could conceive of a constant process of new item writing and cost-effective improvement in functional assessment by improving the items available for individual CAT assessments. This could be accomplished without changing the underlying scoring metric, so that validity data needed to anchor and interpret scores would not have to be redone after every CAT revision.

2. Item Response Theory

IRT methods open the door to understanding the linkages among items used to assess a common outcome domain, and in this way serve as the psychometric foundation underlying CAT. IRT methods examine the associations between individuals' responses to a series of items designed to measure a specific outcome domain (e.g., functional status).[17;33;34] IRT measurement models are a class of statistical procedures used to develop measurement scales. The measurement scales are comprised of items with a known relationship between item responses and positions on an underlying functional domain, called an item characteristic curve. The form of the relationships is typically non-linear. Using this approach, probabilities of patients scoring a particular response on an item at various functional ability levels can be modeled. Persons with more functional ability have higher probabilities of responding positively to functional items than persons with lower functional abilities. These probability estimates are used to determine the individual's most likely position along the functional dimension. When assumptions of a particular item response theory model are met, estimates of a person's functional ability do not strictly depend on a particular fixed set of items. This scaling feature allows one to compare persons along a functional outcome dimension even if they have not completed the identical set of functional items. Since items and functional outcome scores are defined on the same scale,

items can be optimally selected to provide good estimates of each outcome at any level of the scale. This feature of IRT creates important flexibility in administering tests in a dynamic and tailored approach for each individual.

E. THE ACTIVITY MEASURE FOR POST ACUTE CARE

At Boston University, our research group has devoted six years to developing, evaluating, and refining a group of functional status measures built with contemporary measurement techniques. These measures are designed to examine functional status outcomes across the full spectrum of PAC settings including inpatient rehabilitation, long-term care hospitals, skilled nursing facilities, home care, and outpatient rehabilitation clinics. Our main functional status instrument is called the Activity Measure for Post-Acute Care (AM-PAC).[35]

The AM-PAC is a functional status measure developed using the World Health Organization's International Classification of Functioning, Disability and Health (ICF).[36] Within the ICF, an Activity Outcome is defined as "the execution of a task or action by an individual," making it consistent with the definition of functional status used in this chapter. In developing the AM-PAC, we employed two different samples of over 1000 patients who received PAC in acute inpatient rehabilitation units, transitional care units, home care, and ambulatory care settings.[37;38] Based on factor analytic work and IRT analyses,[35;39] three distinct functional status domains have been identified and confirmed:

(a) Physical & Movement Activities;
(b) Personal & Instrumental Activities; and
(c) Applied Cognitive Activities.

The Applied Cognitive Activity domain reflects difficulties or inability to perform basic communication, social and daily management activities (e.g., use of print information, speaking and understanding, interpersonal skills). The domain of Personal Care & Instrumental Activity encompasses primarily basic personal care and instrumental ADLs. The Physical & Movement Activity domain includes basic physical activities such as bending, walking, carrying, or climbing stairs. The classification of specific activities into these three factors is not identical to the ADL and IADL distinction made in Chapter 2 nor Chapter 6; however, these three factors contain a mix of function that may be classified differently by other authors. What they are called is less critical than what is included in each factor.

In developing the AM-PAC, we constructed an underlying functional status item pool of 101 Movement & Physical Activity items, 62 Personal & Instrumental Activity items, and 59 Applied Cognitive Activity items. We developed an initial pool of functional items based on the following: input from measurement and content experts, suggestions solicited from several focus groups of persons with disabilities, and a comprehensive literature review. We selected items underscored by our consumer participants that were perceived as pertinent to their own ability to regain competence in daily activities during a PAC episode. Some items were modified from existing PAC functional instruments such as the FIM, MDS, OASIS, and PF-10, but adapted to the common difficulty and assistance response categories used for AM-PAC item pool. We framed the activity questions in a general fashion without specific attribution to health, medical

conditions, or disabling factors. AM-PAC data were collected by self-report, administered either by the participant's clinician, or by a trained data collector. In 3% of the cases, a proxy respondent (health provider or family member) provided the data.

A series of adaptive short forms and a CAT version of the AM-PAC have been developed and have undergone beta testing in various PAC samples.[40]

We will use a patient case to illustrate below how the AM-PAC CAT works. This case uses the Physical & Movement Activity scale of the AM-PAC. In this scale, we assume that the midpoint of the scale is 50, and this serves as the initial (default) score estimate prior to the CAT administration. For this case example, we used data collected in our Rehabilitation Outcome Study.[41] We set the CAT precision stopping rule as a 95% CI < 3.0. The patient is an individual with congestive heart failure.

Patient A is a 74-year old female who is recovering from congestive heart failure and is being discharged home after an eight-day hospital stay. She has mild arthritis and is leaving the hospital able to only walk for short distances. She is scheduled to have home care visits to assist her with mobility and self-care activities at home. Prior to leaving the hospital, she is asked a series of CAT generated AM-PAC questions regarding her current mobility status. Three months later, she is also asked to respond to CAT-generated AM-PAC questions by her home health care provider to update her functional mobility changes. The results from these two functional assessments are illustrated in Table 5.1.

Table 5.1: Illustrative Functional Assessments Using AM-PAC CAT Methodology

Physical Functioning CAT at Hospital Discharge			**Physical Functioning CAT at Follow-up (3 months post-hospital discharge)**		
Question	Response	Score Estimate and (SE)	Question	Response	Score Estimate and (SE)
Standing up from a chair?	Lot of difficulty	38.4 (7.9)	Standing up from a chair?	No difficulty	44.3 (8.4)
Standing for one minute?	Little difficulty	36.5 (5.4)	Walking outdoors? (100 meters)	No difficulty	56.3 (7.3)
Walking indoors (50 meters)?	Lot of difficulty	35.9 (4.3)	Lifting 10 pound object?	Little difficulty	59.2 (5.3)
On and off toilet?	Little difficulty	36.2 (3.4)	Carrying grocery bag?	No difficulty	62.2 (4.1)
Flight of stairs?	Unable	35.6 (2.9)	Three flights of stairs?	Little difficulty	60.4 (3.0)
Final Hospital Discharge Physical Functioning Score Estimate		**35.6 (2.9)**	**Final 3-month Post-hospital Discharge Physical Functioning Score Estimate**		**60.4 (3.0)**

Note that the patient's responses at hospital discharge indicate that she was functioning at a low level, and therefore the CAT provided AM-PAC questions that addressed these low levels of mobility. At the 3-month follow-up assessment, her responses to the AM-PAC clearly indicated that she was no longer limited in basic mobility and had progressed considerably in physical functioning. The CAT at this phase tailored items in response to the higher levels of functioning

noted by her responses and provided an assessment using more challenging items, yet the two assessments were scored on the same *underlying metric*. During the 3-month period she improved from a 35.6 score to a 60.4 functional level on the AM-PAC Physical & Movement Scale. The CAT provided an estimate of functional ability after she responded to each item, and continued that estimation until some stop-rule based on number of items or precision was satisfied. By adapting to her responses, the CAT yielded questions that were designed especially for her estimated level of ability, and thus provided a precise estimate of her function at each time point with fewer questions than a fixed-length form where she would have been asked the same questions each time.

1. AM-PAC's Sensitivity to Change

Using Adaptive Short Form versions, the utility of the AM-PAC was employed in a Rehabilitation Outcomes Study,[41] to track the functional outcomes of 435 PAC patients.

These adults were recruited at the point of discharge from an acute care hospital or on admission to a rehabilitation hospital and followed for 1, 6, and 12 months through their entire episode of PAC. By linking scores across different Adaptive Short Form versions of the AM-PAC, we tracked the pattern of functional recovery for this cohort across all settings where they received PAC. Sensitivity analyses of the AM-PAC compared with the FIM revealed that the AM-PAC adaptive short forms reduced ceiling effects and increased sensitivity to change at 6 and 12 months follow-ups as compared to the FIM.[42]

In a recently completed project of a 3-month follow-up of patients who were tested after being discharged from inpatient rehabilitation, we found the CAT programs to be equally responsive to patient-reported change over a 3-month interval as compared to fixed-length forms with covering three functional content domains (66 items). On average, the CAT programs required 43% of the time and 33% of the items compared to the fixed-form alternative. We concluded that accurate estimates for group-level functional changes can be obtained from CAT administrations, with a considerable reduction in administration time.[38]

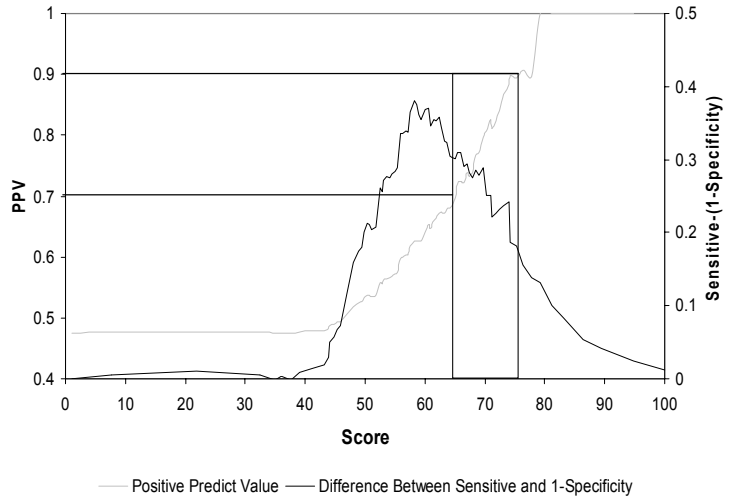
2. Clinical Meaningfulness of AM-PAC Scores

At this stage of development, the clinical meaningfulness of specific scores on each AM-PAC functional domain scale has yet to be identified. However, with additional testing in larger samples of patients receiving PAC, we believe it is feasible to predict clinically important levels of functional status on the AM-PAC scales. As an illustration of how this might be accomplished, we estimated the scale value of the AM-PAC Physical & Movement Activity that optimized the ability to predict unlimited community participation 6 months after a hip fracture for the 53 hip fracture patients in our Rehabilitation Outcomes Study. The Receiver Operating Characteristic Curve (ROC) analysis illustrated in Figure 5.2 graphically depicts the relationship between the Positive Predictive Value (PPV) relative to Sensitivity and False Positive Rate (1-Specificity) across values of the AM-PAC's Physical & Movement Activity scale. The binary outcome being predicted is unlimited community participation measured by answering the following question: "Think about how you go places, using any help or means of transportation available. How much are you limited in getting around?" As the illustration shows, 6 months

after suffering a hip fracture, an AM-PAC Physical & Movement scale score of 70 achieves a PPV of approximately 0.80.

Figure 5.2: Predicting Community Participation with AM-PAC Scores 6 Months Post Hip Fracture

Think about how you go places, using any help or means of transportation available. How much are you limited in getting around?

Some limitation ⓪; Not at all ①	

3. Applications Using the AM-PAC CAT

The AM-PAC has been implemented and/or by a number of private health systems throughout the country, as described below.

a. Merck Phase 2 Clinical Trial (currently underway): The AM-PAC CAT is being used as a functional endpoint in a 24-week, double blinded, randomized, placebo controlled, multi-national, study that is assessing the safety and efficacy of a compound (MK-0677) developed by Merck Pharmaceuticals for the treatment of sarcopenia in patients recovering from hip fracture. In preparation for use as en endpoint in this trial, the AM-PAC was **translated** into 7 languages: British English, Spanish, German, Swedish, Norwegian, Danish, and Hebrew.

b. HealthSouth Outpatient Division (currently underway): During 2005, a AM-PAC CAT Clinical Outcomes Monitoring System was pilot tested with 20 HealthSouth (HS) outpatient rehabilitation practices located in Texas, Connecticut, New Jersey, Florida, and Virginia. During the pilot, 6,000 AM-PAC CAT patient assessments were performed. This pilot was very successful. A new contract was executed in October 2005, to expand from 20 pilot sites to 190 HS sites located in several regions where HS operates. The roll out began in January 2006. HS has identified 450 outpatient sites, which would benefit from the software and the plan is to expand to all sites.

c. SeniorMetrix (currently underway): SeniorMetrix, a case management company, currently uses the FIM instrument as the key functional assessment instrument to assess inpatients. They

are interested in using the AM-PAC as an alternative to the FIM. As a first step, a research study to develop a crosswalk between the two instruments will be jointly undertaken. If the crosswalk is successful, Seniormetrix will begin using the AM-PAC instead of the FIM.

d. Kaiser Permanente of Northern California (currently underway): An agreement was reached with Kaiser to implement the AM-PAC CAT system in Northern California. In 2006, a 6-month pilot study will be conducted in an outpatient site run in the Napa Solano Area. If successful, expansion to other areas within Kaiser is likely. Kaiser has asked to use AM-PAC in a study proposal they are writing in the area of multiple sclerosis.

e. Select Medical (under discussion): Select Medical is a company with over 700 outpatient rehabilitation locations and is similar to HealthSouth. The plan is to establish an AM-PAC CAT monitoring system in their outpatient sites during 2006.

F. DATA COLLECTION METHODS

Understanding variation across respondents who provide information on functional status is an important step in the process of creating a sound outcome instrument for PAC. This examination is especially important if the functional assessment will be used to examine recovery of daily activities over time and across settings, as more than one respondent may be required. Currently, established practice in most hospital and home care settings relies on clinician-based assessments. However, long-term outcome monitoring and examining functional recovery in community settings (e.g., outpatient programs) are likely to require patient-reported data. In addition, some patients cannot provide their own information on function due to limitations in cognitive or communication abilities. Our research has found acceptable proxy-patient reliability (intraclass correlation coefficients (ICC) = 0.68-0.90) across functional domains on the Activity Measure for Post-acute Care (AM-PAC) with the lowest ratings for the items included within the Applied Cognition domain.[43] We have also found that discrepancies are more likely to occur in early stages of recovery during which patients may not be able to successfully self-report.[44] The combination of proxy reports from family members or clinicians with patient report is controversial and information on the variation introduced by proxy respondents is necessary.[45;46] Variance in functional status between respondents would need to be carefully examined and scores would need to be adjusted as systematic effects in respondent differences are established in each functional domain.

G. RECOMMENDATIONS

1. CAT-based Monitoring System

To achieve its long-term goal of longitudinal monitoring of functional status and other relevant outcomes across PAC settings, we propose that CMS consider launching a large-scale demonstration project to examine the feasibility and utility of adopting IRT and CAT methods to develop PAC outcome measures that are psychometrically adequate, comprehensive and precise to monitor change across an entire PAC episode while being practical for widespread application and use across settings. We further recommend that CMS consider using the AM-PAC for the purpose of testing the utility of the IRT/CAT approach to functional status assessment and

monitoring across PAC settings. There are several reasons to consider using the AM-PAC in a field demonstration of this method.

1. The AM-PAC instrument was designed specifically for cross setting functional status monitoring in PAC.
2. IRT item banks for 3 functional domains have been developed and have undergone initial testing and evaluation with over 1000 PAC patients. The item banks have been reviewed by content experts, consumer groups, and have undergone cognitive testing and subsequent revision.
3. Adaptive short form and CAT versions of the 3 AM-PAC functional domains have been developed and have undergone beta testing in PAC patient samples.
4. Given the developmental work already done on the AM-PAC, a large-scale field demonstration could be designed and launched in a short period of time.

Testing of CAT models have been reported in rehabilitation and PAC settings, and early results are promising.[40;47-49] We believe there is sufficient evidence to date to suggest that a CAT system could provide a workable model for longitudinal assessment of functional status outcomes in PAC. Using the CAT method has the potential to decrease the response burden of assessment, while providing an individually-tailored assessment at each recovery stage. The National Institutes of Health have recently included computer-adapted testing approaches as part of their Roadmap,[50] and have funded major multi-year CAT projects for clinical research applications.[51] These major initiatives are designed to develop more uniformity in outcome endpoints used for clinical trials and to capitalize on the efficiency and precision advantages of contemporary approaches towards estimating health status changes. A clear parallel can be drawn for the need to use a similar approach to the measurement requirements of PAC monitoring.

To undertake a field demonstration of the CAT version of the AM-PAC, access to a CAT platform would be needed in each participating hospital, and then a system would need to be in place to access the CAT data in the PAC setting and to complete longitudinal data collection. Providers conducting each assessment would need some type of PDA, stand-alone computer, or Internet access. Point-of-contact computing power is necessary since the order of CAT questions administered requires real time computing power in developing a score estimate and the application of an algorithm to choose an optimal item at after each respondent response. Once established, the CAT platform could provide an optimal system for acquiring accurate and efficient assessments in a uniform PAC outcome monitoring system. If such a CAT platform were not available, an alternative approach would be to use a series of Adaptive Short Form versions of the AM-PAC for the field demonstration.

2. CAT Test Development for Other PAC Outcomes

We recommend that CMS consider undertaking a series of development projects aimed at developing contemporary measurement method approaches to assessing outcomes across PAC settings for outcomes other than functional status. Development of an initial CAT strategy for monitoring other PAC outcomes across PAC settings would take 3-5 years and would require

several stages of methodological development. Let us use the example of cognitive status outcomes again to illustrate each stage of development.

The *first stage* would involve assembling, testing, and revising **item banks** designed to tap each cognitive status domain of interest. For example, one domain could encompass basic reasoning, and another might cover memory or problem solving skills. An item bank consists of individual instrument items or questions that are drawn or written to match a specific functional domain of interest. In the case of cognitive status item banks, items can be drawn from existing fixed-form instruments or newly written. A separate item bank would need to be developed for each domain of cognitive status to be included within the overall instrument. An appropriate rating scale would need to be selected or developed for each cognitive status domain to maximize measurement reliability and validity as well as to maximize respondent understanding.

The second stage would be to conduct cognitive interviews with relevant PAC patient groups to obtain the in-depth information on the patient's understanding of the items included in the item banks and the need for item revision or rewording to enhance understanding and clarity. Cognitive testing can also be useful in identifying gaps in item content within a cognitive domain that may need to be addressed with newly written or revised items.

The third developmental phase would be the application of IRT methodology to evaluate the usefulness of each item included in the item bank, particularly its relative difficulty on the underlying continuum for that particular cognitive domain. This step requires that cognitive item banks be administered to a large sample of PAC patients so that IRT techniques can be subsequently applied to build new instruments, scale responses, and equate scores, as well as to identify item bias that may need to be eliminated. These IRT methods are the foundation on which CAT methodology is implemented. Various IRT mathematical models can be used to estimate the association between an individual's cognitive ability and item difficulty on the instrument's item bank.

In the fourth stage of development work, information on a patient's ability on a particular cognitive domain and the item's known difficulty on the underlying continuum is used to develop efficient CAT algorithms and to equate different forms of a given CAT test with each other. This is the key feature of CAT methodology that allows different users of the same underlying item bank to speak a shared language even though individuals may be responding to different items included in the item bank.

In the fifth and final stage, a series of validation studies would have to be undertaken to provide interpretation guidelines for different levels of cognitive scores generated by the CAT cognitive status instrument. This could take the form of predictive validity studies where different threshold levels of the CAT-generated cognitive status scores are used to predict relevant and important PAC outcomes such as length of stay in a setting, discharge from one care setting to another, or other outcomes of interest such as mortality risk, hospital readmission, or community or social participation levels. A CAT system could ultimately also provide a basis for grouping patients into cognitive groups for use in prospective payment systems.

3. Alternative Strategies

a. Summary Score Cross-Calibration: If CMS is unable to adopt a CAT approach in the short term, a strategy CMS might consider is to equate summary scores from existing functional status instruments using an IRT cross-calibration approach. This method places corresponding items from two instruments on the same scale to develop linkages and summary score correspondence tables.[52;53] Using this IRT approach, one or more functional status instruments could be calibrated to the metric of another instrument, provided the two scales are measuring similar underlying functional constructs. This provision is critical in using an IRT approach, since a key assumption of these models is that items from one or more instruments are one-dimensional.

The summary score cross-calibration approach would require a calibration sample of patients who completed each instrument that was to be cross-calibrated. A series of IRT analyses would be completed to confirm the unidimensionality of the combined item set, an estimation of the item parameters, and the development of expected value curves superimposed on the combined underlying metric.

By way of illustration, Figure 5.3 displays a cross-calibration we developed of scores from the six MDS-PAC functional items and five mobility items of the FIM (unpublished data). Using item parameters developed across the two instruments, we developed expected scores for each instrument at each level of the overall mobility continuum. The FIM items are on a 5-35 metric, and the MDS-PAC is on a 0-35 metric. On both scales, higher scores indicate increasing physical functioning ability. An expected total score can be obtained for either the AM-PAC or the FIM at any level of the underlying scale. Based on this figure, a correspondence table (Table 5.2) could be developed that links the scores from one instrument to another.

Figure 5.3: Cross-calibration of Scores from Six MDS-PAC Functional Items and Five FIM Mobility Items

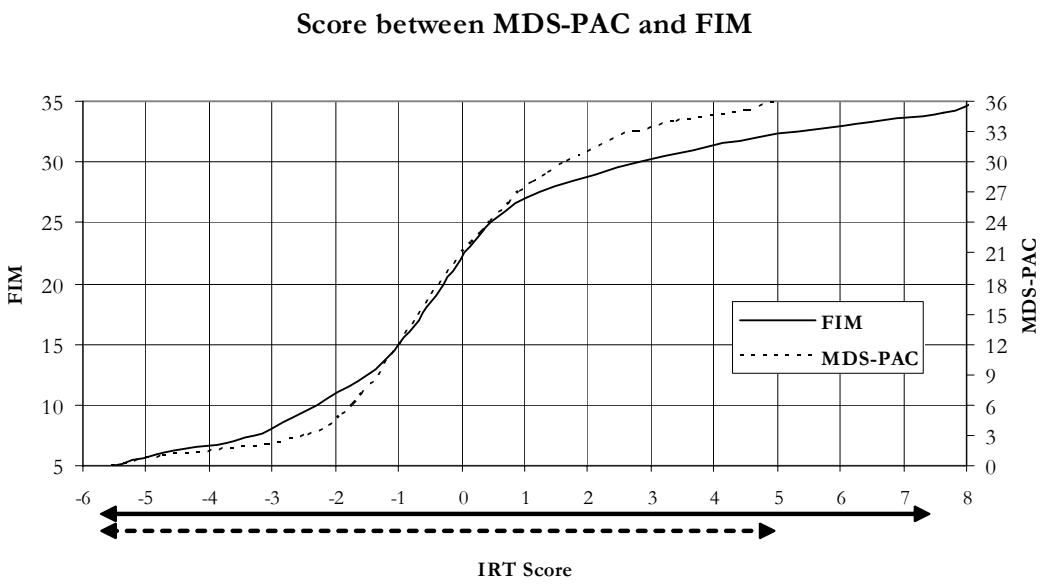


Table 5.2: Score Correspondence Between FIM Mobility Items and MDS-PAC Mobility Items

Underlying Scale of Mobility Functioning (logits)	FIM Mobility Items (raw scores)	MDS-PAC Mobility Items (raw scores)
-4.0	7	2
-2.0	11	5
0	22	21
1.0	29	31
4.0	32	34

A final step in this process would be to confirm the validity of the expected scores to the observed scores from a new cross-validation sample.

An advantage of this approach is that it combines items into a common metric. However, if the content coverage of the items within the two instruments is similar, then no additional content coverage beyond the range of either instrument becomes available. If the items from the two instruments cover different content, then a broader range of functional status can be assessed by a combination of items from the two instruments, yet placed on a similar scale. Note in Figure 5.3 that the FIM has a broader content range than the MDS-PAC at the higher range of mobility functioning. This, of course, is the important advantage of using IRT methods; by combining items from multiple instruments, one can develop a uniform scale with broader content coverage than with any one instrument alone.

b. Develop Adaptive (Parallel) Short-forms: A final interim strategy that CMS might consider is the development and application of Adaptive Short Form instruments. There are several examples in health care applications that use adaptive short forms for measuring a functional construct across different levels of expected patient functioning[54] or between PAC settings.[37] The adaptive short form approach uses IRT as the basis for linking the two forms together on a common metric. This technique has the advantage of providing a means to compare functional scores across persons using different yet parallel short forms. Using adaptive short forms for different groups of patients increases scaling efficiency over a single fixed form applied to everyone. Respondents need only respond to one form (a subset of the functional items that best targets their level of function based on the setting in which they are receiving PAC). While CAT scales would yield the largest increases in efficiency,[40] adaptive short forms can provide an initial step toward improving measurement across PAC settings while not increasing response burden. A key characteristic of the adaptive short form approach is that the functional items selected and used in the different short forms must all come from the same item bank. This approach links the forms together so that scores from each form can be compared, avoiding the problems of combining scores from two instruments that have not been co-calibrated using IRT on the same metric.

We recently developed a series of adaptive short forms to measure physical functioning across the PAC continuum.[37] We developed a short form of 10 optimal items for application in a

hospital or facility setting, and 12 optimal items for patients who had returned to a community setting. Our reasoning was that questions on physical functioning for patients who were in the hospital should be centered toward the low ability end, and should not include outdoor walking or other items that would be meaningless to assess in the hospital setting. For community assessment (patients receiving home care or outpatient services), questions that were selected for inclusion were broader and evaluated higher levels of home and community mobility. The items selected for both forms were from an established AM-PAC item bank, and so the items were linked to an underlying continuum and common score metric.[35] In a recent field trial of these adaptive short-forms in a 12-month follow-up of patients receiving rehabilitation, the adaptive forms were more sensitive to long-term changes at a 6- and 12-month follow-up period in comparison to a standard fixed form used typically at an early recovery stage in inpatient rehabilitation.[42]

Consider the following case example:
Patient B is a 66-year old male who is recovering from a fractured hip and pelvis from an automobile accident. He is being discharged home post-surgery after a six-day acute care hospital stay. He is only partially weight-beating on affected hip when discharged from the acute care hospital. He is scheduled to have home care visits to assist with mobility and transfer activities at home. Prior to leaving the hospital, he is asked a series of questions from a 10-item survey on his current physical functioning abilities. At six months after discharge, he is sent a form that asks him to update his current physical functioning status.

Note in Table 5.3 that the two adaptive forms have different questions, which are related to expected physical functioning in either the hospital or home setting. The ability to use the two parallel forms helps to spread the range of functioning that can be covered by questionnaires with relatively small numbers of items.

Table 5.3: Illustrative Functional Assessments Using Adaptive Short Form Methodology

Physical Functioning Adaptive Short Form at Acute Care Hospital Discharge		**Physical Functioning Adaptive Short Form for Home/ Community Settings (6-months post-hospital discharge)**	
Question	Response	Question	Response
Sitting up in bed?	Lot of difficulty	Sitting up in bed?	No difficulty
Reach in standing?	Unable	Walking outdoors? (100 meters)	Little difficulty
Chair transfer?	Unable	Pick object off floor?	Little difficulty
Bending or kneeling?	Unable	One-mile walk?	Unable
Bed mobility?	Little difficulty	Flights of stairs?	Little difficulty
Use bathroom?	Total assistance	Walk within home?	No difficulty
Bed transfer?	Total assistance	Walk indoors (50 meters)?	Little difficulty
Walk in hospital room?	Total assistance	Moderate activities?	Lot of difficulty
Walk in hospital hallway?	Total assistance	Vigorous activities?	Unable
Flight of stairs	Total assistance	Walk outdoors several blocks?	Lot of difficulty
Hospital Discharge Physical Functioning Score Estimate = 20.6		**6-month Post-hospital Discharge Physical Functioning Score Estimate= 55.4**	

Although use of parallel short forms improves the efficiency and precision of measurement over a single fixed-form, this approach has not been shown to be as effective in estimating scores as the CAT, especially for persons at the low and high end of functional ability.[40]

H. CONCLUSIONS

Precise measurement of functional status at all levels is a critical feature of any outcome assessment system proposed for quality of care determination across the entire continuum of PAC. ***Computer Adaptive Testing (CATs)*** coupled with ***Item Response Theory (IRT)*** provides a way to achieve practical and precise measurement and tracking of relevant outcomes throughout an entire episode of PAC. Although relatively new in the health field, CAT/IRT methodology is already commonplace in other fields. As an initial step toward developing a PAC monitoring system, we propose that CMS launch a demonstration project to examine the feasibility and utility of adopting IRT and CAT functional status instruments to monitor functional change across an entire PAC episode of care using the Activity Measure for Post Acute Care, developed by researchers at Boston University. Testing of CAT models has been reported in rehabilitation and PAC settings, and early results are promising. We believe there is sufficient evidence to suggest that a CAT system could provide a workable model for longitudinal assessment of functional status outcomes in PAC. Using the CAT method has the potential to decrease the response burden of assessment, while providing an individually-tailored assessment at each recovery stage throughout an episode of PAC.

I. REFERENCES

(1) Liu K, Gage B, Harvell J, Stevenson D, Brennan N. Medicare's post-acute care benefits: background, trends, and issues to be faced. 1999.

(2) Gage B. The balanced budget act: implication for post-acute care services. 1998.

(3) Testimony of Herb Kuhn, Director, Center for Medicare Management, Centers for Medicare Medicaid Services. 6-16-2005. US House Ways and Means Subcommittee on Health, Hearing on Post Acute Care.

(4) Johnson M, Holthaus D, Harvell J, Coleman E, Eilertsen T, Kramer A. A Medicare post-acute care quality measurement: final report. Report prepared for the Office of the Asst Sec for Planning and Eval, editor. 2001. University of Colorado Health Sciences Center.

(5) National Committee on Vital and Health Statistics: Classifying and Reporting Functional Status. 2002.

(6) Shaughnessy PW, Hittle DF, Crisler KS, Powell MC, Richard AA, Kramer AM et al. Improving patient outcomes of home health care: findings from two demonstration trials of outcome-based quality improvement. Journal of the American Geriatrics Society 2002; 50:1354-1364.

(7) Clark MS, Bond MJ, Prior KN, Cotton AC. Measuring disability with parsimony: evidence for the utility of a single item. Disability and Rehabilitation 2004; 26:272-279.

(8) Sloan JA, Aaronson N, Cappelleri JC, Fairclough DL, Varricchio C, Clinical Significance Consensus Meeting Group. Assessing the clinical significance of single items relative to summated scores. Mayo Clinic Proceedings 2002; 77(5):479-487.

(9) Harvey R, Jellinek HM. Functional performance assessment: a program approach. Archives of Physical Medicine and Rehabilitation 1981; 62(9):456-461.

(10) Shaughnessy P, Crisler KS, Schlenker RE. Medicare's OASIS: Standardized Outcome and Assessment Information Set for Home Health Care - OASIS B. Center for Health Services and Policy Research. Denver, CO: 1997.

(11) Ware J, Sherbourne C. The MOS 36-item short form health survey (SF-36): conceptual framework and item selection. Medical Care 1992; 30(6):473-483.

(12) Guide for the Uniform Data Set for Medical Rehabilitation (including the FIMTM instrument), Version 5.1. 1997. Buffalo, NY, State University of New York at Buffalo.

(13) Minimum Data Set (MDS) for Nursing Home Resident Assessment and Care Screening. 1998.

(14) Ness K. The use of the MDS 2.0 to measure rehabilitation outcomes in skilled nursing facilites. Journal of Rehabilitation Outcomes Measurement 2000; 4(2):22-30.

(15) Fortinsky RH, Garcia RI, Sheehan J, Madigan EA, Tullai-McGuinness S. Measuring disability in Medicare home care patients: application of Rasch modeling to the outcome and assessment information set. Medical Care 2003; 41(5):601-615.

(16) Nunnally J. Psychometric Theory. New York, NY: McGraw-Hill Book Company, 1967.

(17) Hambleton RK. Emergence of item response modeling in instrument development and data analysis. Medical Care 2000; 38(9):II-60.

(18) Buchanan JL, Andres PL, Haley SM, Paddock SM, Zaslavsky AM. An assessment tool translation study. Health Care Financing Review 2003; 24(3):45-60.

(19) Buchanan JL, Andres PL, Haley SM, Paddock SM, Zaslavsky AM. Evaluating the planned substitution of the Minimum Data Set-Post Acute Care for use in the rehabilitation hospital prospective payment system. Medical Care 2004; 42:155-163.

(20) Weiss D. Improving measurement quality and efficiency with adaptive testing. Applied Psychological Measurement 1982; 6:473-492.

(21) Velozo CA, Kielhofner G, Lai JS. The use of Rasch analysis to produce scale-free measurement of functional ability. American Journal of Occupational Therapy 1999; 53(1):83-90.

(22) McHorney CA. Generic health measurement: past accomplishments and a measurement paradigm for the 21st century. Annals of Internal Medicine 1997; 127(8):743-750.

(23) Jette A, Haley SM, Ni P. Comparison of functional status tools used in post-acute care. Health Care Financing Review 2003; 24(3):13-24.

(24) Morris J, Murphy K, Nonemaker S. Long-term Resident Care Assessment User's Manual, Version 2.0. 1995.

(25) Ware J, Kosinski M. The SF-36® Physical and Mental Health Summary Scales: A Manual for User's of Version 1.0. 2nd ed. Boston, MA: Health Assessment Lab, 2001.

(26) Chen A, Broadhead W, Doe E, W B. Patient acceptance of two health status measures: the Medical Outcomes Study Short-Form General Health Survey and the Duke Health Profile. Family Medicine 1993; 25(8):536-539.

(27) Haley S, Jette AM. RRTC for measuring rehabilitation outcomes: extending the frontier of rehabilitation outcome measurement and research. Journal of Rehabilitation Outcomes Measurement 2000; 4:31-41.

(28) Cook K, Monahan P, McHorney C. Delicate balance between theory and practice: health status assessment and item response theory. Medical Care 2003; 41(5):571-574.

(29) Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. Medical Care 2000; 38(9s):II-28-II-42.

(30) Revicki DA, Cella DF. Health status assessment for the twenty-first century: item response theory, item banking and computer adaptive testing. Quality of Life Research 1997; 6:595-600.

(31) Wainer H. Computerized Adaptive Testing: A Primer. New Jersey: Lawrence Erlbaum Associates, 2000.

(32) Jette AM, Haley S. Contemporary measurement techniques for rehabilitation outcomes assessment. Journal of Rehabilitation Medicine 2005; 37(6):339-345.

(33) Embretson SE, Reise SP. Item Response Theory for Psychologists. Mahwah, NJ: Lawrence Earlbaum Associates, 2000.

(34) Thissen D. Data analysis using Item Response Theory. Psychological Bulletin 1988; 104:385-395.

(35) Haley SM, Coster WJ, Andres PL, Ludlow LH, Ni PS, Bond TLY et al. Activity outcome measurement for post-acute care. Medical Care 2004; 42(Suppl 1):I-49-I-61.

(36) World Health Organization. ICF: International Classification of Functioning, Disability and Health. Geneva: World Health Organization, 2001.

(37) Haley SM, Andres PL, Coster WJ, Kosinski M, Ni PS, Jette A. Short-form activity measure for post-acute care (AM-PAC). Archives of Physical Medicine & Rehabilitation 2004; 85:649-660.

(38) Haley S, Ni P, Hambleton R, Slavin M, Jette A. Computer adaptive testing improves accuracy and precision of scores over random item selection in a physical functioning item bank. Journal of Clinical Epidemiology (In press) 2006.

(39) Coster WJ, Haley SM, Andres PL, Ludlow LH, Bond TLY, Ni PS. Refining the conceptual basis for rehabilitation outcome measurement: personal care and instrumental activities domain. Medical Care 2004; 42(Suppl 1):I-62.

(40) Haley SM, Coster WJ, Andres PL, Kosinski M, Ni PS. Score comparability of short-forms and computerized adaptive testing: simulation study with the Activity Measure for Post-Acute Care (AM-PAC). Archives of Physical Medicine and Rehabilitation 2004; 85:661-666.

(41) Jette A, Keysor J, Coster W, Ni PS, Haley SM. Beyond function: predicting participation outcomes in a rehabilitation cohort. Archives of Physical Medicine and Rehabilitation 2005; 86(11):2087-2094.

(42) Coster W, Haley S, Jette A. Measuring patient-reported outcomes after discharge from inpatient rehabilitation settings. Journal of Rehabilitation Medicine (In press) 2006.

(43) Andres PL, Haley SM, Ni PS. Is patient-reported function reliable for monitoring post-acute outcomes? American Journal of Physical Medicine & Rehabilitation 2003; 82(8):614-621.

(44) Haley S, Ni P, Coster W, Black-Schaffer R, Siebens H, Tao W. Agreement in functional assessment: graphical approaches to displaying respondent effects. American Journal of Physical Medicine & Rehabilitation (In press) 2006.

(45) Magaziner J, Simonsick EM, Kashner MT, Hebel RJ. Patient-proxy response comparability on measures of patient health and functional status. Journal of Clinical Epidemiology 1988; 41(11):1065-1074.

(46) Rothman ML, Hedrick SC, Bulcroft KA, Hickam DH, Rubenstein LZ. The validity of proxy-generated scores as measures of patient health status. Medical Care 1991; 29(2):115-124.

(47) Dijkers MP. A computer adaptive testing simulation applied to the FIM instrument motor component. Archives of Physical Medicine and Rehabilitation 2003; 84:384-393.

(48) Andres PL, Black-Schaffer RM, Ni PS, Haley SM. Computer adaptive testing: a strategy for monitoring stroke rehabilitation across settings. Top Stroke Rehabil 2004; 11(2):33-39.

(49) Siebens H, Andres P, Ni P, Coster W, Haley S. Measuring physical function in patients with complex medical and postsurgical conditions: a computer adaptive approach. American Journal of Physical Medicine & Rehabilitation 2005; 84(10):741-748.

(50) National Institute of Health. The NIH Roadmap: Re-engineering the Clinical Research Enterprise. Department of Health and Human Services.

(51) Fries J, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. Clinical and Experimental Rheumatology 2005; 23(5 (suppl 39)):S53-S57.

(52) Linn RL. Linking results of distinct assessments. Applied Measurement in Education 1993; 6:83-102.

(53) Cook LL, Eignor DR. Using item response theory in test score equating. International Journal of Educational Research 1989; 13:161-173.

(54) Cook KF, Roddey TS, Gartsman GM, Olson SL. Development and psychometric evaluation of the Flexlevel Scale of Shoulder Function. Medical Care 2003; 41(7):823-835.

Chapter 6
Conclusions and Recommendations

In this chapter, we begin in Section A with a summary of the major findings in the report, which are referenced in earlier chapters. Section B discusses the recommended domains and justification for these domains in a uniform assessment instrument. Section C presents a long-term vision and recommendations for assessment in post-acute care (PAC). Section D proposes a set of short-term recommendations, with particular emphasis on the first-year activities that would lead to the long-term vision.

A. SUMMARY OF REPORT FINDINGS

1. Purposes of the Proposed Uniform Assessment Instrument

This 10-week project was conducted to examine existing approaches to assessment for PAC and make recommendations to CMS on how to proceed with respect to uniform assessment. In PAC, we are including the inpatient settings of skilled nursing facilities (SNFs), inpatient rehabilitation facilities (IRFs), and acute long-term care hospitals (ALTCs), as well as residential-based care provided by home health agencies (HHAs) and outpatient programs. The three purposes of the instrument that we have highlighted for this report include: 1) making placement decisions related to the most appropriate PAC setting and services at hospital discharge; 2) serving as a core set of information that should be transmitted to the receiving provider to enhance the safety and quality of care transitions; and 3) providing baseline information for longitudinal follow-up of health and function, elements of which would be repeated over time. Given the settings that need to be covered and these three purposes, in essence the uniform assessment instrument needs to suit all Medicare beneficiaries discharged from the hospital.

The major tradeoff in developing such an instrument is balancing burden with the desired set of information to characterize the full spectrum of the population of interest. For example, some might argue that disease-specific measures for the major diagnoses ought to be incorporated into the instrument. However, the factors that are most important for these three purposes generally cut across conditions, and while disease-specific information is critical for care planning, it is not essential for these purposes. Thus, the assumption for this uniform assessment instrument is that it would be completed at hospital discharge, providing information that is useful to the receiving PAC site, but would not substitute for a clinical assessment conducted at the site directed toward care planning and treatment. If CMS ultimately gravitates toward comparable site-specific assessments, this uniform assessment instrument could underlie the site-specific assessment tools, reducing the assessment burden at the PAC sites.

Identifying uniform assessment items for the purposes of determining payment or resource needs was also beyond the scope of this project. The proposed uniform assessment instrument is not intended to replace the payment variables that are included in the existing instruments. However, the fact that the proposed domains will be important for placement determinations and identification of transitional care needs suggests that some of the most important payment items will be included (although not necessarily using the same metrics that are used in the current

instruments). For example, payment rates in the RUGs, CMGs, and HHRGs are driven more by measures of physical function than any other domains, with the possible exception of service variables such as therapy provided. While the uniform assessment instrument may not include the exact same functional scales as used in those tools, function will most certainly be included, and the proposed metric could ultimately be used to recalibrate payments across settings.

A final point is that the purposes of an instrument and the population of interest drive the selection of both the appropriate domains and specific measures. Approaching the development of the uniform assessment instrument without initially clarifying the purposes and population would provide no framework for including or excluding specific domains and for deciding on the appropriate measures. Thus, although it may seem easiest to just select an assessment tool that was developed for some specific purpose within a certain population, such an approach would not lead to the optimal cross-site assessment tool with the lowest possible burden.

2. Current Post-Acute Care Instruments

The MDS, OASIS, and IRF-PAI were developed from very different historical roots, without the intent of uniformity across assessments. In addition, they were developed for different purposes, none of which pertain to the first two purposes of the proposed instrument. As a result, lack of consistency in domains, measures, time intervals, and the nature of the questions makes it extremely difficult to generate clear and comparable metrics across these instruments. None of these instruments, nor the vast majority of the measures in these instruments, can serve as the basis for a uniform assessment tool. However, work is underway to cross-walk or map the instruments to uniform metrics.

A major issue in using one of these instruments across settings or rendering the instruments comparable is that the items are not sensitive to conditions or functional abilities of patients across the full spectrum of the PAC population. For example, the MDS functional items are targeted at assessing function in nursing homes, and in fact were originally developed for the long-term nursing home population, which is generally more dependent than the IRF or HHA populations. The IRF-PAI functional items (FIM) were designed to be sensitive to changes in the IRF population, which is typically less dependent than the SNF population and more dependent than the home health population; whereas the OASIS items were designed for the less impaired home health population. In addition, the SF-36 was designed for the ambulatory care population with functional items that are sensitive to these patients' status. Thus, none of the scales are sensitive to functional status across the full spectrum of function, and the difference in the items and time frames makes it very difficult to scale them using the same metric.

A major limitation of the existing systems is that they use outdated questionnaire and measurement designs, paper forms, and an infrastructure that once put in place are extremely difficult to change. The data collection and measurement methods used result in extremely burdensome forms because the same set of items is asked of every individual, even when a substantial portion of the items are irrelevant for a specific person and could be avoided based on previous responses. This situation would be exacerbated if the uniform assessment instrument used the same questionnaire design because the spectrum of patients across PAC settings is far broader than the spectrum in any given setting. In addition, until we move away from the

widespread use of questionnaires administered on paper forms that have to be entered and transmitted, we not only are expending excessive resources on this process but also will continue to be locked into a rigid set of questions. Thus, the existing tools do not take advantage of information technologies or questionnaire design and measurement development techniques that now exist.

3. Non-CMS Instruments and Measures

Several attempts have been made to develop uniform assessment instruments for various purposes: the VA Geriatrics and Extended Care (GEC) Referral Form, the Health Outcomes Survey (HOS), the Uniform Needs Assessment Instrument (UNAI), and the Continuing Care Record (CCR). Our conclusion after carefully reviewing these is that we do not recommend that any of them be used in its entirety by CMS as the uniform assessment instrument. These instruments generally focus on one segment of the population (e.g., HOS is largely for outpatient care) or have such gross measures of impairment that precise measurement of impairment across the full PAC spectrum would not be feasible (e.g., the VA GEC form has 9 dichotomous ADL/physical mobility items). Thus, the items would not meet all the purposes for the populations of interest for Medicare PAC. They also have some of the other weaknesses of existing CMS PAC instruments involving a fixed item set requiring complex infrastructure, making the questions difficult to revise.

Nevertheless, these initiatives all offer useful information both for measures of specific domains that will be necessary for the uniform assessment instrument, and technologies relating to IT infrastructure and large-scale data acquisition and monitoring efforts. Measures that have been developed, tested, and/or are in use, which will ultimately be recommended for the uniform assessment can and should be proposed from these various tools. Through consensus and testing, these efforts have certainly built knowledge regarding the most important domains, as well as measures for those domains. In addition, many measures exist for the various domains of function, cognition, depression, etc. that have been developed in either comprehensive assessment tools or specific research activities. Development of a uniform assessment instrument should fully consider this research base in selecting measures rather than starting anew wherever possible.

4. Placement and Care Transitions

Although most would argue that patient and caregiver factors should drive placement decisions for PAC, other system and reimbursement issues such as bed availability, geographic area, and availability of coinsurance or co-payments, come into play. Evidence-based criteria or algorithms for optimal placement in the most cost-effective settings for most patients with different conditions and characteristics are lacking for PAC; however, the information set that is important to consider in making placement decisions has been suggested by a number of studies, which we will present within the discussion of recommended domains. Lack of a uniform information set at hospital discharge has made it impossible to develop specific norms for different settings, much less algorithms for clinicians to use as guidelines for decision-making. The importance of patient- and caregiver-reported items in this information set cannot be underestimated.

Providing a core set of information to patients, families, and receiving clinicians following hospital discharge has become increasingly evident to ensure safe and efficient care transitions. From the hospital side, this activity falls under the broad rubric of discharge planning, which is being more fully incorporated into the JCAHO methodology for hospital quality review. A number of studies have demonstrated the benefits of communicating an accurate core set of information accompanied by interventions that assure this information is utilized. These studies have identified important domains for care transitions. A critical lesson from these studies is the need for assessment information that is both timely and accurate. Ideally, such information would be available electronically through interoperable records, the web, or email. Alternatively, fax or sending a paper copy with the beneficiary has also been used successfully.

5. Health Outcome and Quality Assessment

Several domains can be identified in which baseline data and longitudinal follow-up are important for monitoring outcome and quality of different PAC settings. Functional outcomes are the most widely applicable and complex domains for PAC. In this report, we discuss four major domains of function including physical mobility, ADLs/personal care, IADLs/advanced cognitive activities, and social functioning. For monitoring outcomes in each of these domains, we require metrics that can accommodate the full spectrum of patients at baseline and are sensitive to changes in these metrics across this spectrum of patients as they improve or decline.

Item Response Theory (IRT) and Computer Adaptive Technology (CAT) are well developed methods for minimizing the number of items that need to be asked of a given respondent, while mapping all of these items to the same metric (see Chapter 5). For mobility, for example, there may be 50 items in the item pool, but based on each answer, the respondent is directed to an appropriate question, and inappropriate and unnecessary questions are never asked. So a respondent who cannot transfer out of bed is never asked questions about walking a block and the person who is able to walk a mile is never asked questions about whether he or she can move around in bed. This dramatically reduces the burden of the instrument, maps all items on the same metric regardless of where an individual falls on the scale, and is most sensitive to change for each individual because questions are targeted in the area where he or she falls on the scale. Another advantage is that new items can be added and tested, and items can be refined while retaining the same metric.

These methods are not only developed for applications such as national educational testing, but have been used for functional measurement and shown to meet these objectives. The Activity Measure for Post-Acute Care (AM-PAC), a functional status measure developed by researchers at Boston University over the last six years, is currently in use by providers such as HealthSouth, Kaiser Permanente of Northern California, and used for measurement in trials by Merck. Other provider entities such as Senior Metrix are considering using this metric if it can be mapped adequately to the FIM. Collection of information using this methodology is optimized with the use of a PC or web application in which all the logic is programmed. Testing has also suggested that patient response measures (in which a proxy is used if the patient cannot respond) are reliable when using these methods for functional assessment. The SF36 and the HOS as well as

numerous other tools have relied on patient response in order to obtain truly patient-centered measures of outcome.

A final consideration of health outcome assessment is the need for collection of these uniform data at fixed intervals regardless of the admission and discharge pattern among the individual PAC providers. PAC episodes frequently involve multiple providers with relatively short stays in each provider, making it necessary to assess outcomes using a fixed interval following hospital discharge regardless of location. These outcome measures therefore would not replace outcome monitoring within settings conducted by individual facilities for the purposes of clinical assessment and quality improvement, but will allow comparison across different types of PAC episodes. The appropriate body to oversee longitudinal data collection for PAC is not clear, but just as NCQA serves this function for the Medicare HOS data, a central agency could be responsible.

B. RECOMMENDED DOMAINS

The purpose of this section is to provide initial recommendations on domains for a uniform assessment instrument that serves the three purposes previously discussed. The three purposes of discharge placement, care transitions, and monitoring outcomes/quality do not necessarily require the same set of information, but the uniform assessment instrument must include domains that are necessary for any one of these purposes.

Potential measures and measurement issues for these domains have been discussed throughout the report. In this section, we do not recommend specific measures because agreement on the set of domains is an essential first step. However, key measurement issues for the different domains are highlighted in the text. This enumeration of domains is based on the literature and previous consensus activities in these areas that were uncovered during this 10-week project. The evidence base, however, is not sufficient to defend each of these domains based on research, nor is it sufficient to exclude all other possible domains. Thus, this should be seen as a recommendation that would benefit from more widespread confirmation.

Table 6.1 provides an enumeration of the recommended domains, indicating which of the three purposes for which they will be used and a recommended respondent for each domain. Respondents are categorized either as "P" denoting the patient or a proxy for the patient (which may be a family member or other caregiver who is involved with the daily care of an individual), or "H" denoting the item could be populated by a hospital database or by a hospital staff person from their records. The assumption is that the information will be completed in the hospital as near to hospital discharge as possible, with the exception of the outcome/quality measures that would also be completed at fixed intervals in addition to baseline. The justification for each domain is summarized below Table 6.1 based on the information contained in this report.

Table 6.1: Recommended Domains for the Three Purposes of the Uniform Assessment Instrument

Domains	Discharge Placement	Care Transitions	Outcomes/ Quality	Respondent*
Goals of care	X	X		P
Specialized rehab care needs	X			H
Patient's residence & with whom	X		X	P
Meets Medicare criteria for homebound	X			H
Active problem list	X	X		H
Medication list	X	X		H
Allergies/intolerances		X		H
Resuscitation status/advance directive/DPAHC		X		P
Discharge instructions/outstanding diagnostic tests		X		H
Cognitive functional status	X	X		P
Physical functioning/mobility	X	X	X	P
Activities of Daily Living/self-care	X	X	X	P
IADLs/Advanced cognitive	X	X	X	P
Social functioning	X		X	P
Premorbid Function	X	X		P
Self-rated health status		X	X	P
Pain status		X		P
Depression	X	X		P
Skin integrity		X		H
Sensory deficits	X	X		P
Dietary needs		X		H
Continence	X	X		P
Fall risk		X		H
Services receiving in home	X	X		H
DME receiving/equipment	X	X		P, H
Able and willing caregiver	X	X		P
Ethnic or cultural considerations/language	X	X		P
Immunizations and most recent PPD test result		X		H
Infectious precautions	X			H
Insurance/financial resources	X	X		P
Basic demographics; age, gender	X	X		H

*P= Patient/Proxy; H=Populated by hospital database or hospital staff

1. Goals of Care: The patient's goals of care, which should be clarified with the assistance of hospital staff, are required for appropriate discharge placement and to prepare the receiving provider for expectations. Goals may include, but are not limited to, rehabilitation, skilled nursing, prevention of adverse events such as hospitalization, palliation or end of life care, and stabilization of certain conditions. More than one goal may be appropriate.

2. Specialized Care Needs: An enumeration of specialized care needs is necessary in determining the appropriate discharge location, including services such as PT, OT, ST, oxygen, tracheostomy care, ventilators, ostomy care, tube feeding, IV infusions, dialysis, wound care, etc. These needs can factor heavily into determinations of appropriate PAC placement and care as discussed in Chapter 2.

3. Patient's Residence and With Whom: For discharge placement decisions, the patient's place of residence just prior to the hospitalization is critical information. A long-term nursing home resident will more than likely return to the nursing home even following a period of PAC; therefore, PAC services must be oriented toward return to the nursing home (which in fact might mean using a rehabilitation program in the same nursing home in which the resident had been residing in order to minimize the number of transitions). In addition to discharge placement, many studies have used the outcome of residential location as an indicator of rehabilitation effectiveness. Although a gross indicator of functional recovery, return to a comparable setting enhances quality of life, particularly for individuals previously residing in the community.

4. Meets Medicare Criteria for Homebound: Meeting the Medicare homebound criteria is necessary for discharge to a HHA and thus constitutes essential information for discharge placement. According to Medicare, the criteria include normal inability to leave home except with considerable and taxing effort, and absences from home are infrequent or of relatively short duration or are attributable to receiving medical treatment. Individuals meeting the homebound criteria are not required to receive PAC in an HHA.

5. Active Problem List: An active problem list, including a patient's current diagnoses and health care problems, is necessary information for receiving providers and to determine the setting to which the patient should be discharged. Most of this information can be obtained directly from the hospital record, but uniformity is required as discussed in Chapter 2.

6. Medication List: As discussed in Chapter 4, miscommunication about medications lead to care transition failures, resulting in rehospitalizations and other misadventures. Thus, accurate medication lists including dosages and administration should be included in the uniform assessment, which could potentially be populated from the hospital database. To the extent possible, the medication list should be reconciled with the admitting medications, indicating whether the medications are new or have been suspended, and/or any altered medication dosages or instructions.

7. Allergies/Intolerances: Similar to medications, allergies and intolerances should be recorded to ensure such information is available to receiving providers. Although providers

will typically ask the same questions, this critical information may be available only in records for individuals who are cognitively impaired.

8. Resuscitation Status/Advance Directive/Durable Power of Attorney for Health Care: Although this information will most likely be confirmed at the receiving site, the hospital information on resuscitation and advanced directives should be included. If the patient has advance directives or durable power of attorney for health care, the receiving provider can rely on that information if the person is cognitively impaired or unable to communicate.

9. Discharge Instructions/Follow-Up/Diagnostic Tests and Appointments: Clear discharge instructions and scheduled follow-up diagnostic tests and appointments are necessary for any Medicare beneficiary leaving the hospital. Otherwise, there is serious risk that he/she will encounter problems following hospitalization, may not know who to contact, and may soon end up back at the hospital or emergency department.

10. Cognitive Functional Status: For placement decision-making and to prepare the receiving provider, a specific measure of cognitive function is important (Chapter 3). Of most importance is that the measure covers the spectrum of cognitive functioning that Medicare beneficiaries being discharged from the hospital might experience, while the precision of the estimate may be less critical at this stage (except at certain key decision points). For example, placement in an IRF depends on individuals being able to follow fairly complex instructions, such as two-step commands; otherwise, motor learning will be difficult.

11. Physical Functioning/Mobility: This domain includes the full spectrum of mobility from moving in bed to walking several miles. This is critical to placement and care transitions as seen in Chapter 4 and an essential outcome measure as seen in Chapter 5. Both covering the spectrum and a reasonable degree of precision in this measurement are necessary in order to use the measure for outcome monitoring.

12. Activities of Daily Living/Self-Care: This domain includes the basic activities of at least toileting, dressing, bathing, personal hygiene, and eating. Because these activities are required daily for personal care, the information is essential for discharge placement, to prepare receiving providers, and as an important outcome domain distinct from physical functioning/mobility. One problem is that while most beneficiaries being discharged from the hospital may be independent in all of these activities, the scale needs to be very sensitive to small differences for those who are dependent because such differences can have a strong impact on where these beneficiaries can reside.

13. IADL/Advanced Cognitive Function: This domain includes a wide array of activities required to live in the community including such tasks as housekeeping, transportation, managing money, shopping, taking medications, using the telephone, and various aspects of communication. While different items have been included in this domain by different authors, the general construct is recognized as distinct from the more basic personal care activities. These activities all require levels of both cognitive and often physical abilities that are more complex. Once again, this scale must be sensitive to gradations of

measurement across the entire scale because individuals may range from being extremely independent in all of these activities to being virtually unable to accomplish any of them.

14. Social Functioning: Several studies show that social functioning, including participation in recreational activities and interaction with others, is important for discharge planning and placement. This functional domain is less well developed than the others that have been measured over many decades in PAC and long-term care. However, recent studies suggest the importance of social/role functioning as a marker for quality of life outcomes.

15. Premorbid Function: Functional status prior to the hospital event has been found in several studies to be an important consideration in discharge placement. A patient's recovery is clearly limited by his or her function prior to an acute event, which needs to be considered in the context of goals and prognosis.

16. Self-Rated Health Status: Self-perceived health status provides a global measure for providers about a patient's quality of life. Poorly rated health status is associated with high utilization of health care services, which should alert providers to individuals who are most vulnerable during care transitions. As an outcome measure, self-rated health status is an important domain that should benefit from effective PAC.

17. Pain Status: Pain has been called the fifth vital sign because it is so important to a patient's health and wellbeing, and it is treatable. Several measures exist for pain (Chapter 3), and hospitals focus on it increasingly. Thus, a rating of pain status and a plan for pain control are necessary at care transition.

18. Depression: A depression screen is important to detect the potential for depression that would then require further evaluation and treatment following discharge. As discussed in Chapter 3, without a screen, depression is under-reported, which can seriously hamper care.

19. Skin Integrity: Frequently a patient is discharged to a PAC setting with skin problems such as pressure sores, stasis ulcers, surgical wounds, or abrasions. Because it is essential to begin monitoring and managing skin problems within hours of admission to a PAC setting, including an assessment of skin is necessary for care transitions.

20. Sensory Deficits: Receiving providers can benefit immensely from information gleaned in the hospital through testing visual, hearing, and other sensory deficits. These deficits are typically thoroughly evaluated in the hospital and are important to consider in making placement decisions and preparing receiving providers to avoid unsuccessful transitions.

21. Dietary Needs: To the extent that the hospital has assessed an individual's nutrition and dietary needs, this information should be made available to the receiving provider. This is particularly important if dietary needs have changed as a result of the hospitalization and a patient may not fully understand the implications of the new diet.

22. Continence: Urinary or bowel incontinence may mean the difference between whether an individual can receive PAC in the home or in a nursing home, depending upon the

availability of assistance. Furthermore, receiving providers should not have to discover past problems with continence because new continence problems may arise due to the stress of the transition or some other factors.

23. Fall Risk: A simple assessment of fall risk could prevent a major fall and fracture in the period immediately after care transition because the receiving environment is not appropriately fall proof. Whether this is a discharge to home or to a facility, knowledge of fall risk may decrease an untoward event.

24. Services Receiving in the Home: Whether a patient has any paid help in the home or is receiving volunteer services could affect discharge placement. However, proper communication is necessary during the care transition in order to assure that these services are available and are still appropriate and sufficient given the individual's new status.

25. Durable Medical Equipment (DME)/Equipment: Both previous and newly ordered DME needs to be specified. This could influence discharge placement and is necessary in the care transition. DME such as special beds and mattresses, ambulation equipment (e.g., trapeze, walker, cane, wheel chair), or specialized equipment for ADLs will be necessary as soon as an individual arrives at a PAC setting.

26. Able and Willing Caregiver: Availability of an able and willing caregiver has been found to strongly predict PAC placement and can most certainly make a difference as to whether an individual can go home initially or even over the long run if the disability is likely to persist. Both whether the caregiver is able to provide the necessary assistance and whether the caregiver is willing must be assessed based on a patient's condition. Because patients have a range of needs and caregivers have a range of abilities, this domain must capture a broad spectrum of circumstances.

27. Ethnic or Cultural Considerations/Language: Frequently, cultural differences will alter the likelihood that family will be interested in caring for a chronically ill or dying person in the home. Language issues should also be taken into consideration in planning discharge and assuring that care transitions are successful. Communication and cultural understanding may be key to the success of PAC.

28. Immunizations and PPD Test: The hospital discharge provides one more point of care to assure that the appropriate immunizations and a PPD test have been completed.

29. Infectious Precautions: For immune compromised patients, or patients with contagious or high-risk infections, precautions must be known and considered in discharge placement.

30. Insurance/Financial Resources: A patient's financial resources need to be taken into consideration even when Medicare is the primary payor because of co-payments, coinsurance, and the fact that some services or options may not be covered by Medicare. At very least, patients need to be aware of the costs of different alternatives.

31. Basic Demographics – Age, Gender: Age and gender are necessary both to consider placement from the perspective of a patient's capacity for intensive rehabilitation and to recover and live independently. They are also important identifiers in the care transition.

C. LONG-TERM VISION AND RECOMMENDATIONS

The long-term recommendations are intended to provide a vision for assessment across PAC settings that could be achievable over the next five to ten years depending upon commitment to such a strategy. The reason for providing this long-term vision is that any short-term recommendations should feed into a longer-term plan or we risk continuing to develop setting-specific assessments along paths that will never meet. Many details of this long-term recommendation have not been refined, but the vision is what is important.

1. Core Dataset Completed at Every Hospital Discharge

A core dataset with the necessary domains for placement decisions, transitions, baseline outcome assessment, and ultimately payment for episodes of PAC should be completed for every Medicare beneficiary upon hospital discharge. Questions for all domains would be embedded in a uniform instrument completed on an automated platform. Current measurement strategies such as IRT/CAT should be used in order to minimize respondent burden while maximizing precision of measurement for each beneficiary. While these methodologies are not applicable to all domains, particularly those not generally assessed on a continuum (e.g., place of residence, medications, care plans, etc), they should be used for relevant domains that can be scaled on an underlying metric (e.g., physical function, and social function). Items would consist of both provider response and patient/proxy interview to assure patient centeredness. Provider-report information would be populated to the extent possible from existing hospital systems using standardized vocabularies, minimizing burden and data entry error.

The product generated from the assessment would be a summary report (electronic and hardcopy) providing information such as the medication list, advance directives, allergies, and the metrics for the functional items and other scales for domains such as cognition and depression. The scores could ultimately be converted to percentiles for the population of discharges, which would be more meaningful to users who are unfamiliar with each specific metric. Thus, the uniform assessment would be distilled to a summary report (probably one page) without inclusion of responses to each and every item that were asked of the individual beneficiary.

2. Algorithms Developed for Recommended Patient Placement

Based on research using the core database, norms for characteristics of patients treated most cost effectively in the different settings could be determined. To the extent possible, algorithms or data-driven recommendations on placement based on these norms could be generated from the dataset. Placement decision guidance could be overridden based on unique characteristics of the beneficiary, the community, the caregiver, or bed availability as appropriate.

3. Uniform Assessment Would Be Transmitted Electronically to Receiving Providers

The summary obtained from the assessment conducted at hospital discharge with all the key information would be transmitted directly to the receiving provider, preferably through access to a protected website. Alternatively, email or even use of interoperable records would be possible. The transmitted information would assure continuity of care from the hospital without replacing an assessment conducted by the site for clinical purposes. However, the core dataset could feed directly into standard assessment tools used by each type of PAC setting so that patients would continue to be monitored on the same set of metrics. For IRT-based metrics, greater precision in the scoring of individuals using these metrics would be possible at sites by the use of more questions from the item pool to refine the individual score. This would enable providers to detect smaller but significant changes in some of these metrics as part of care activities and quality improvement. The other domains for any federally-mandated standard PAC assessment would not be duplicative of those in this uniform assessment tool, but would build upon it in areas pertinent to specific PAC providers.

4. Health and Outcome Monitoring System

The hospital discharge assessment would provide baseline data for longitudinal follow-up of key functional outcomes, quality of life, residence, and utilization. This information would be collected at fixed intervals regardless of PAC settings and lengths of stay. These longitudinal outcome data would be utilized in a database to determine expected outcomes for comparable patients across settings in order to make future policy decisions about appropriate coverage.

5. Payment Would be Based on Similar Metrics Regardless of Discharge Setting

One option would be to use data generated from the uniform assessment and longitudinal cost measurement to develop payment models for total episodes of PAC. These could be used to make per episode payments that covered the entire multi-provider episode rather than setting-specific payments. This strategy would encourage the use of cost-effective settings and avoid the incentive to shorten stays and discharge to subsequent providers, resulting in more multiple-provider episodes. Alternatively, use of the same metrics for functional status would be possible in each setting-specific payment system in order to equate patient groups. Thus, setting-specific payment systems could be developed based on similar items.

D. SHORT-TERM RECOMMENDATIONS

The following set of recommendations provides a possible road map for developing and testing a uniform assessment instrument that could then be implemented for national demonstration activities. These recommendations assume that the project would build upon the findings of this report with a CMS endorsement such that an extensive national vetting of all decisions would not be required through town meetings and national consensus activities. Obviously, any assessment instrument would need to be sufficiently defensible to withstand critique, but engagement of all parties in the development would likely derail efforts to produce an instrument focused on specific purposes that minimizes burden to providers and patients.

1. Review of Recommended Domains for Uniform Assessment Instrument

A group of technical experts with expertise in all assessment settings, the Medicare program, and research should review the set of domains proposed in this report for the three purposes. The technical expert panel should be sent this report along with a table similar to Table 6.1 that includes literature citations and a definition for each domain. Panel members should be asked to rate each domain in terms of relative importance prior to the meeting; the ratings should be compiled and then discussed at the meeting. Other possible domains should be discussed as well. Although one should expect topics such as specific measures and measurement strategies to arise during the discussion, the focus of this activity should be on the domains. We have found that consensus on domains when the purpose and population is clear is not that difficult to achieve among unbiased experts.

2. Testing Functional Measurement Using IRT/CAT

Using the AM-PAC platform developed by Boston University researchers, a test could be conducted using specified IRT-based functional domains and any relatively straightforward items pertaining to identified non-IRT-based domains (e.g., demographics, advanced directives, diagnoses). The intent of the testing would be to have some users begin to collect hospital discharge data related to function and make it available to receiving providers. The item pool developed for the IRT could be enhanced as needed to meet the needs of the acute hospital discharge population. Preliminary data could help generate norms for different PAC providers.

3. Identify or Develop Measures for Each Domain

To the extent that tested measures exist, these should be identified for each domain and rated in terms of the extent to which they cover the PAC spectrum, validity, reliability, and burden. A combination of provider- and patient-response questions should be utilized depending upon the nature of the domain. Provider response information pertaining to medications, ICD-9 codes and allergies, for example, should be structured using standardized vocabularies, increasing the likelihood that such fields can be populated directly from electronic health records and can be transported into records at the receiving provider. Patient response items should be worded as they would be asked of patients, and algorithms for selecting proxies should be included in the instrument logic. Cross-walking to existing CMS PAC assessment tools and the VA GEC Referral Form should be considered in item selection.

Although an attempt should be made to use measures and items from existing tools that have been tested and meet criteria for validity and reliability, when no such instrument exists that applies directly to this purpose, modest developmental efforts will be necessary. These efforts should begin by review and recommendation of experts in that specific domain in order to take advantage of what is currently in use and under development, but may not be published. Targeted testing on small samples to refine questions and metrics should then be completed before incorporating into the larger instrument. Screening questions and skip patterns should be considered to reduce respondent burden when items do not apply. Because the instrument will initially be set up as an automated tool, items can be tested and refined as development proceeds.

4. Automate the Tool

For administering, maintaining, and transmitting a uniform assessment instrument, a web application is the ideal approach. A website can be accessed with sufficient firewalls for security purposes by the hospital staff person completing the assessment, who would enter information and would be led through an interview by the application involving specific questions for the patient. Range checks can be included in the software, no data entry is necessary, and the information could be entered at different times if at one point the resident interview was completed and at another point the provider information was completed or uploaded. The web-based application would generally be easy to access from any PC on the Internet and would be easy to update without disseminating new software to individual locations. However, a PC-based version of the tool and logic could readily be developed as well, the results of which could be uploaded or transmitted.

The assessment information could be retrieved via the website by receiving providers who have system access. The information could also be emailed or downloaded by the site, as appropriate. Although electronic transmission always raises concerns about security, the system could be completely HIPAA compliant and through encoding the data stream could be completely secure (as in banking). The issues do not arise with security of the data stream; rather, they relate to restricting user access and avoiding inadvertent sharing of access codes and passwords.

Other hardware/software platforms currently under development or available to CMS should also be explored.

5. Beta Testing

Testing in a sample of hospitals for a range of Medicare PAC beneficiaries would be necessary for refining questions and the metrics and item pools for functional domains. A balance between precision and burden would have to be determined in all domains. Software and the technology would be refined through this testing. This beta testing would also give initial norms for the various metrics that could prove useful for demonstration design. It could be relatively simple to enroll a large number of patients very fast in this type of test because it could be used for any Medicare beneficiary discharged from the hospital. An important policy question, however, is the incentive for hospitals to complete such a form as part of the testing. Without some type of incentive it will be difficult to motivate discharge planners or other hospital staff to complete the tool.

As part of the testing, longitudinal follow-up at fixed intervals for the portion of the instrument that would be used for outcome monitoring should be conducted. Presumably this follow-up would be by telephone with the beneficiary or caregiver, but telephone follow-up could be compared with in-person follow-up during this initial test. The difference between proxy and patient responses to items should also be tested during this period in order to understand biases that might occur for any of the outcome items. This outcome follow-up should be conducted at several intervals to determine the most appropriate intervals for different types of PAC.

6. National Demonstration

Following pilot testing and refinement, a national demonstration could then be conducted, yielding a large database of uniform assessments and longitudinal data. The information could be used to: 1) compare patients discharged to different types of settings or multiple-provider episodes and determine the extent to which there is substitution; 2) compare costs for comparable patients across different episodes of care; 3) develop placement algorithms for clinician guidance and determining the most appropriate post-acute care settings; and 4) assess the effects of the transmitted uniform assessment instrument on care transitions. Ultimately, such information would prove useful for working toward a more safe, cost-effective, patient-centered system of PAC for Medicare beneficiaries.