NATIONAL QUALITY FORUM
National Voluntary Consensus Standards for Nursing Homes 2010

Measure Number/Title: NH-022-10: Percent of Residents Whose Need for Help with Activities of Daily Living Has Increased (Long Stay)

Description: This measure is based on data from the MDS 3.0 assessment of long-stay nursing facility residents and reports the percentage of all long-stay residents in a nursing facility whose need for help with late-loss Activities of Daily Living (ADLs), as reported in the target quarter's assessment, increased when compared with a previous assessment. The four late-loss ADLs are: bed mobility, transferring, eating, and toileting. This measure is calculated by comparing the change in each item between the target MDS assessment (which may be an annual, quarterly or significant change or correction assessment) and a previous assessment (which may be an admission, annual, quarterly or significant change or correction assessment).

Numerator Statement: The numerator is the number of long-stay residents who have an MDS assessment (which may be an annual, quarterly, significant change, or significant correction) reporting a defined amount of decline when compared with a previous assessment (which may be an admission, annual, quarterly, significant change, or significant correction MDS 3.0 assessment). This would indicate an increase, when compared with a previous assessment, in the resident's need for help with a late-loss item as indicated by a higher score (coding convention is such that a higher score indicates the need for more help with a task). The need for increased assistance (suggesting decline in function) is identified if the score for at least one late-loss ADL item increases by two or more points or if the score for two or more of the late-loss ADLs items increase by one point; late-loss ADL items are bed mobility, transferring, eating, and toileting.

Denominator Statement: The denominator includes all long-stay residents who received an annual, quarterly or significant change or correction MDS 3.0 assessment during the quarter and who did not meet the exclusion criteria.

Level of Analysis: Population: national, Facility/Agency

Data Source: Electronic clinical data

Measure developer: Research Triangle Institute International

Type of Endorsement (full or time-limited): Full

Attachments: Activities of Daily Living Table

NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized): Complete all <mark>yellow highlighted</mark> areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: If there is no TAP or workgroup, the SC also evaluates the subcriteria (<mark>yellow</mark> highlighted areas).

Steering Committee: Complete all <mark>pink</mark> highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** NH-022-10   **NQF Project:** Nursing Homes 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 Measure Title: Percent of Residents Whose Need for Help with Activities of Daily Living Has Increased (Long Stay)
De.2 Brief description of measure: This measure is based on data from the MDS 3.0 assessment of long-stay nursing facility residents and reports the percentage of all long-stay residents in a nursing facility whose need for help with late-loss Activities of Daily Living (ADLs), as reported in the target quarter's assessment, increased when compared with a previous assessment. The four late-loss ADLs are: bed mobility, transferring, eating, and toileting. This measure is calculated by comparing the change in each item between the target MDS assessment (which may be an annual, quarterly or significant change or correction assessment) and a previous assessment (which may be an admission, annual, quarterly or significant change or correction assessment).
1.1-2 Type of Measure: Outcome **De.3 If included in a composite or paired with another measure, please identify composite or paired measure**
De.4 National Priority Partners Priority Area: Safety **De.5 IOM Quality Domain:** Safety **De.6 Consumer Care Need:**

CONDITIONS FOR CONSIDERATION BY NQF	NQF Staff
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* **A.1 Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes **A.2 Indicate if** Proprietary Measure (*as defined in measure steward agreement*):	A Y ☐ N ☐

A.3 Measure Steward Agreement: Government entity and in the public domain - no agreement necessary **A.4** Measure Steward Agreement attached:	
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	B Y ☐ N ☐
C. The intended use of the measure includes <u>both</u> public reporting <u>and</u> quality improvement. ▶**Purpose:** Public reporting, Internal quality improvement	C Y ☐ N ☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. **D.1** Testing: Yes, fully developed and tested **D.2** Have NQF-endorsed measures been reviewed to identify if there are similar or related measures? Yes	D Y ☐ N ☐
(for NQF staff use) Have all conditions for consideration been met? Staff Notes to Steward (*if submission returned*):	Met Y ☐ N ☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*): Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. *Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	Ev al Rat ing
(for NQF staff use) Specific NPP goal:	
1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers, Patient/societal consequences of poor quality **1a.2** **1a.3 Summary of Evidence of High Impact:** Increasing ADL dependence is associated with decreased quality of life. Greater dependency has been shown to be a risk factor for complications such as pressure ulcers, hospitalizations, and reduced quality of life.(1, 2) Although some ADL decline may be unavoidable resulting from circumstances of the individual's clinical conditions, ADL deterioration can also result from inadequate nursing care or rehabilitation therapies.(3) Risk factors for functional decline include injuries, medication side effects, pain, poor nutrition, the use of restraints, prolonged bed rest, and the prolonged use of indwelling catheters.(1, 2) These factors may be mitigated by nursing care, multidisciplinary communication, and referrals for rehabilitation therapies and nutrition services. In addition, improved physical environmental factors (e.g., chairs with arms, improved lighting) may contribute to maintaining or improving function.(2) ADL decline is also associated with substantial Medicare costs. In a study focused on a community-residing sample, 10.0% beneficiaries who declined in function accounted for more than 20.0% of hospital, outpatient, and nursing facility expenditures.(6)	1a C ☐ P ☐ M ☐ N ☐

Comment [KP1]: 1a. The measure focus addresses:
•a specific national health goal/priority identified by NQF's National Priorities Partners; OR
•a demonstrated high impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use (current and/or future), severity of illness, and patient/societal consequences of poor quality).

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

Using MDS 2.0 data for April–June 2008, the national prevalence of ADL decline in nursing facilities was 16.1%, with a range of 10.6% in Oregon to an average of 24.2% in North Dakota. The national measure results have been stable over time, ranging from 15.4% in 2002 to 14.9% in 2008.(6)

1a.4 Citations for Evidence of High Impact: 1. Colorado Foundation for Medical Care. Development, maintenance, and implementation of nursing home quality measures. Environmental scan: Review of the literature, clinical guidelines, and other sources for information pertinent to the CMS publicly reported nursing home quality measures. Final draft working team document with abstracts. Englewood, CO: Colorado Foundation for Medical Care, 2007.

2. Centers for Medicare & Medicaid Services. MDS Quality measure/indicator report. 2008. Available from http://www.cms.hhs.gov/MDSPubQIandResRep/02_qmreport.asp?isSubmitted=qm3&group=05&qtr=14.

3. Degenholtz HB, Rosen J, Castle N, Mittal V, Liu D. The association between changes in health status and nursing home resident quality of life. Gerontologist. 2008 Oct;48(5):584-92.

4. Horn SD, Buerhause P, Bergstrom N, Smout RJ. RN staffing time and outcomes of long-stay nursing home residents: Pressure ulcers and other adverse outcomes are less likely as RNs spend more time on direct patient care. Am J Nurs. 2005;105(11):58-70.

5. Kresevic DM. Evidence-based geriatric protocols for best practice. 3rd ed. New York: Springer Publishing Company, Inc., 2008.

6. Fried T, Bradley E, Williams C, Tinetti M. Functional disability and health care expenditures for older persons. MD Arch Intern Med. 2001;161:2602-7.

7. American Health Care Association. Trends in publicly reported nursing facility quality measures. 2008. Available from http://www.ahcancal.org/research_data/trends_statistics/Documents/trends_nursing_facilities_quality_measures.pdf.

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: This measure is intended to prompt nursing facilities to evaluate whether their long stay residents are experiencing avoidable ADL declines and if so, develop approaches to help their residents improve or maintain their function. The benefit envisioned by use of this measure is improved functional status in long stay nursing facility residents and concomitant improvements in residents' quality of life.

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
This quality measure has been used by CMS since 2002. Except for minor clarifications in the definitions and instructions, the ADL measures have not changed in the transition from MDS 2.0 to MDS 3.0.

In the University of Colorado's analysis of the current quality measure on ADL decline using MDS data from 2006 (as shown in the following table), the rate of ADL decline varied substantially across facilities. The quality measure varied from 5.9% at the 10th percentile to 27.0% at the 90th percentile. Only 0.9% of facilities had no residents with a decline in late-loss ADL function. Thus there appears to be a sizable performance gap across facilities.

See attached Table 1: Measure Variability Across Facilities.

1b.3 Citations for data on performance gap:
1. Brega A, Hittle D, Goodrich G, Kramer A, Conway K, Levy C. Empirical review of publicly reported nursing home quality measures. Denver: Division of Health Care Policy and Research University of Colorado at Denver; Abt Associates, Inc, 2007.

1b
C ☐
P ☐
M
☐
N
☐

Comment [KP2]: 1b. Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating considerable variation, or overall poor performance, in the quality of care across providers and/or population groups (disparities in care).

Comment [k3]: 1 Examples of data on opportunity for improvement include, but are not limited to: prior studies, epidemiologic data, measure data from pilot testing or implementation. If data are not available, the measure focus is systematically assessed (e.g., expert panel rating) and judged to be a quality problem.

1b.4 Summary of Data on disparities by population group:
Although research suggests racial disparities in the quality of care in nursing facilities between African-Americans and Caucasians (1, 2, 3, 4, 5), no studies show evidence of disparities in ADL decline by race, once other demographic factors were taken into account. After controlling for income and education, which were lower among African-Americans, there were no additional declines between African-Americans than Caucasians.(6) In a study of ADL decline in five states, neither gender nor race influenced the trajectory of ADL dependence.(7)

Research has shown that the risk of ADL decline increases with 1) increasing cognitive impairment among nursing facility residents (e.g., 7, 8, 9), although it is unknown the extent to which this reflects poor care rather than unavoidable decline, and 2) with the initiation of dialysis among nursing home residents with end-stage renal disease for whom additional efforts are recommended to maintain ADL function.(10)

1b.5 Citations for data on Disparities:
1. Smith D, Feng Z, Fennell M, Zinn J, Mor V. Separate and unequal: racial segregation and disparities in quality across U.S. nursing homes. Health Aff (Millwood). 2007;26(5):1448-1558.

2. Howard D, Sloane P, Zimmerman S, Eckert J, Walsh J, Buie V, Taylor P, Koch G. Distribution of African Americans in residential care/assisted living and nursing homes: more evidence of racial disparity? Am J Public Health. 2002;92(8):1272-7.

3. Grabowski D. The admission of blacks to high-deficiency nursing homes. Med Care. 2004;42(5):456-64.

4. Mor V, Zinn J, Angelelli J, Teno J, Miller S. Driven to tiers: Socioeconomic and racial disparities in the quality of nursing home care. Milbank Q. 2004;82(2):227-56.

5. Miller SC, Papandonatos G, Fennell M, Mor V. Facility and county effects on racial differences in nursing home quality indicators. Soc Sci Med. 2006;63(12):3046-59.

6. Peek CW, Coward RT, Henretta JC, Duncan RP, Dougherty MC. Differences by race in the decline of health over time. Journal of Gerontology: Social Sciences. 1997;52B(6):S336-S344.

7. McConnell ES, Pieper CF, Sloane RJ, Branch LG. Effects of cognitive performance on change in physical function in long-stay nursing home residents. J Gerontol Med Sci. 2002;57A(12):M778-84.

8. Chen JH, Chan D, Kiely DK, Morris JN, Mitchell SL. Terminal trajectories of functional decline in the long-term care setting. J Gerontol Med Sci. 2007;62A(5):M531-6.

9. Ang YH, Su SYL, Yap LKP, Ee CH. Functional decline of the elderly in a nursing home. Singapore Med J. 2006;47(3):219-24.

10. Tamura MK, Covinsky KE, Chertow G, Yaffe K, Landefeld CS, McCulloch,CE. Functional status of elderly adults before and after initiation of dialysis. N Engl J Med. 2009;361:1539-47.

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*): ADL decline among nursing facility residents is an important health outcome. Impaired functional status is a major cause of institutionalization and facility residents are at risk for further decline, making this measure a central outcome of interest among this population.

1c.2-3. Type of Evidence: Evidence-based guideline, Expert opinion

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*):
The evidence supporting the relationship between nursing care and ADL decline is summarized in Evidence-

1c
C ☐
P ☐
M ☐
N ☐

Comment [k4]: 1c. The measure focus is:
• an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed;
OR
• if an intermediate outcome, process, structure, etc., there is **evidence** that supports the specific measure focus as follows:
o Intermediate outcome – evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
o Process – evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
o Structure – evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
o Patient experience – evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
o Access – evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
o Efficiency – demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

Comment [k5]: 4 Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multi-step process, the step with the greatest effect on the desired outcome should be selected as the focus of measurement. For example, although assessment of immunization status and recommending immunization are necessary steps, they are not sufficient to achieve the desired impact on health status – patients must be vaccinated to achieve immunity. This does not preclude consideration of measures of preventive screening interventions where there is a strong link with desired outcomes (e.g., mammography) or measures for multiple care processes that affect a single outcome.

Based Geriatric Protocols for Best Practice (3rd edition), which also provides guidelines about specific interventions. Risk factors for functional decline include injuries, medication side effects, pain, poor nutrition, the use of restraints, prolonged bed rest, and the prolonged use of indwelling catheters. These factors may be mitigated by nursing care, multidisciplinary communication, and referrals for rehabilitation therapies and nutrition services. In addition, improved physical environmental factors (e.g., chairs with arms, improved lighting), exercise, and socialization may contribute to maintaining or improving function.(1) Although this evidence-based clinical guideline targets care of hospitalized elders to prevent ADL decline, the same evidence applies to the ability of nurses and other facility staff to maximize physical function and prevent or minimize decline in long-term care settings.

1. Kresevic DM. Evidence-based geriatric protocols for best practice. 3rd ed. New York: Springer Publishing Company, Inc., 2008.

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom*):
The body of evidence supporting this measure has not been rated. Kresevic et al rated the strength of the individual studies used in developing the clinical guidelines, which ranged from Level II – Level VI using the scale described below.

1c.6 Method for rating evidence: The individual studies cited were rated by Kresevic et al (2008) as ranging from Level II – Level VI using the following definitions:

Level I: Systematic reviews (integrative/meta-analyses/clinical practice guidelines based on systematic reviews)
Level II: Single experimental study (RCTs)
Level III: Quasi-experimental studies
Level IV: Non-experimental studies
Level V: Care report/program evaluation/narrative literature reviews
Level VI: Opinions of respected authorities/consensus panels

1c.7 Summary of Controversy/Contradictory Evidence: This comment is based on a specific study that evaluated the association between the ADL QM and Medicaid payment policy. In addition, the RUGS classification system used by over 30 states to set their Medicaid payments, is based substantially on levels of ADL impairment. The items underlying the other QMs either do not contribute to the RUGS or do not contribute substantially (e.g., incontinence and pressure ulcers).

There is consensus regarding the importance and centrality of evaluating a facility's ability to minimize or prevent resident ADL decline. The current measure of ADLability in the MDS 3.0 reflects resident need for staff support and is not a self-performance assessment. An increase in ADL score results from the need for an increase in staff oversight or support and reflects an increase in dependency. While some degree of decline may be unavoidable from circumstances resulting from the individual's clinical condition, the expected trajectory is unknown. Indeed, some researchers have assumed the decline shown using MDS data reflects the natural course of decline.(1, 2, 3). Case mix factors may also contribute to the extent of decline observed in a facility (i.e., cognitive impairment).(2, 3, 4, 5) In addition, there is some evidence that ADL decline reported in the MDS is sensitive to Medicaid payment policies (i.e., more ADL decline is reported in states that incorporate this information into their payment formulae [6]), suggesting that state-level policy differences account for some of the observed decline, perhaps through providing an incentive to record decline. Finally, there is disagreement about the reliability of the ADL items upon which the measure is based. Although comparisons between gold-standard nurses had high kappas, other analyses have shown discrepancies in the ADL ratings.(7) To address this, the Rand Corporation developed and tested new ADL measures for inclusion in the MDS, under contract to CMS.(8) However, to avoid undue burden to states using these measures in their payment formulae, CMS postponed incorporating these new items into the MDS and provided some clarifications in the RAI Manual to improve reliability.

1. Peek CW, Coward RT, Henretta JC, Duncan RP, Dougherty MC. Differences by race in the decline of health over time. J Gerontol Soc Sci. 1997;52B(6):S336-44.

2. McConnell ES, Pieper CF, Sloane RJ, Branch LG. Effects of cognitive performance on change in physical function in long-stay nursing home residents. J Gerontol Med Sci. 2002;57A(12):M778-84.

Comment [k6]: 3 The strength of the body of evidence for the specific measure focus should be systematically assessed and rated (e.g., USPSTF grading system http://www.ahrq.gov/clinic/uspstf07/methods/benefit.htm). If the USPSTF grading system was not used, the grading system is explained including how it relates to the USPSTF grades or why it does not. However, evidence is not limited to quantitative studies and the best type of evidence depends upon the question being studied (e.g., randomized controlled trials appropriate for studying drug efficacy are not well suited for complex system changes). When qualitative studies are used, appropriate qualitative research criteria are used to judge the strength of the evidence.

3. Chen JH, Chan D, Kiely DK, Morris JN, Mitchell SL. Terminal trajectories of functional decline in the long-term care setting. J Gerontol Med Sci. 2007;62A(5):M531-6.

4. Phillips C, Shen S, Chen M, Sherman M. Evaluating nursing home performance indicators: An illustration exploring the impact of facilities on ADL change. Gerontologist. 2007;47(5);683-9.

5. Ang YH, Su SYL, Yap LKP, Ee CH. Functional decline of the elderly in a nursing home. Singapore Med J. 2006;47(3):219-24.

6. Bellows NM, Halpin HA. MDS-based state Medicaid reimbursement and the ADL-decline quality indicator. Gerontologist. 2008;48(3):324-29.

7. Abt Associates, Inc.; Stepwise Systems, Inc.; Qualidigm. Data Assessment and Verification (DAVE 2) project—MDS two-stage discrepancy findings, April–December 2006. Cambridge, MA: Abt Associates, Inc., 2007.

8. Saliba D, Buchanan J. Development and validation of a revised nursing home assessment tool: MDS 3.0. Contract No. 500-00-0027/Task Order #2. Santa Monica, CA: Rand Corporation, Apr 2008. Available from http://www.cms.hhs.gov/NursingHomeQualityInits/Downloads/MDS30FinalReport.pdf.

1c.8 Citations for Evidence (*other than guidelines*): 1. Bates-Jensen BM, Alessi CA, Cadogan M, Levy-Storms L, Jorge J, Yoshii J, Al-Samarrai NR, Schnelle JF. The minimum data set bedfast quality indicator: Differences among nursing homes. Nurs Res. 2004 Jul-Aug;53(4):260-72.
2. Conn VS, Minor MA, Burks KJ, Rantz MJ, Pomeroy SH. Integrative review of physical activity intervention research with aging adults. J Am Geriatr Soc. 2003 Aug;51(8):1159-68. [30 references]
3. Counsell SR, Holder CM, Liebenauer LL, Palmer RM, Fortinsky RH, Kresevic DM, Quinn LM, Allen KR, Covinsky KE, Landefeld CS. Effects of a multicomponent intervention on functional outcomes and process of care in hospitalized older patients: a randomized controlled trial of acute care for elders (ACE) in a community hospital. J Am Geriatr Soc. 2000 Dec;48(12):1572-81.
4. Covinsky KE, Palmer RM, Kresevic DM, Kahana E, Counsell SR, Fortinsky RH, Landefeld CS. Improving functional outcomes in older patients: Lessons from an acute care for elders unit. Jt Comm J Qual Improv. 1998 Feb;24(2):63-76.
5. Cunningham GO, Michael YL. Concepts guiding the study of the impact of the built environment on physical activity for older adults: A review of the literature. Am J Health Promot. 2004 Jul-Aug;18(6):435-43. [48 references]
6. Edington J, Barnes R, Bryan F, Dupree E, Frost G, Hickson M, Lancaster J, Mongia S, Smith J, Torrance A, West R, Pang F, Coles SJ. A prospective randomised controlled trial of nutritional supplementation in malnourished elderly in the community: clinical and health economic outcomes. Clin Nutr. 2004 Apr;23(2):195-204.
7. Engberg S, Sereika SM, McDowell BJ, Weber E, Brodak I. Effectiveness of prompted voiding in treating urinary incontinence in cognitively impaired homebound older adults. J Wound Ostomy Continence Nurs 2002 Sep;29(5):252-65.
8. Graf C. Functional decline in hospitalized older adults. Am J Nurs. 2006 Jan;106(1):58-67, quiz 67-8. [33 references]
9. Hirsch CH, Sommers L, Olsen A, Mullen L, Winograd CH. The natural history of functional morbidity in hospitalized older patients. J Am Geriatr Soc. 1990 Dec;38(12):1296-303. [15 references]
10. Hodgkinson B, Evans D, Wood J. Maintaining oral hydration in older adults: a systematic review. Int J Nurs Pract. 2003 Jun;9(3):S19-28. [32 references]
11. Inouye SK, Rushing JT, Foreman MD, Palmer RM, Pompei P. Does delirium contribute to poor hospital outcomes? A three-site epidemiologic study. J Gen Intern Med. 1998 Apr;13(4):234-42.
12. Kresevic D, Holder C. Interdisciplinary care. Clin Geriatr Med. 1998 Nov;14(4):787-98. [12 references]
13. Kresevic DM, Counsell SR, Covinsky K, Palmer R, Landefeld CS, Holder C, Beeler J. A patient-centered model of acute care for elders. Nurs Clin North Am. 1998 Sep;33(3):515-27. [6 references]
14. Landefeld CS, Palmer RM, Kresevic DM, Fortinsky RH, Kowal J. A randomized trial of care in a hospital medical unit especially designed to improve the functional outcomes of acutely ill older patients. N Engl J Med. 1995 May 18;332(20):1338-44.
15. Pedersen BK, Saltin B. Evidence for prescribing exercise as therapy in chronic disease. Scand J Med Sci

Sports. 2006 Feb;16(Suppl 1):3-63. [735 references]
16. Sager MA, Rudberg MA. Functional decline associated with hospitalization for acute illness. Clin Geriatr Med. 1998 Nov;14(4):669-79.
17. Tucker D, Molsberger SC, Clark A. Walking for wellness: A collaborative program to maintain mobility in hospitalized older adults. Geriatr Nurs. 2004 Jul-Aug;25(4):242-5.
18. Vass M, Avlund K, Lauridsen J, Hendriksen C. Feasible model for prevention of functional decline in older people: Municipality-randomized, controlled trial. J Am Geriatr Soc. 2005 Apr;53(4):563-8.

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number*):
Assessment recommendation #5: Interdisciplinary/multidisciplinary collaboration

Management recommendation #1: Maximization of function and prevention of decline

1c.10 Clinical Practice Guideline Citation: Assessment of function in Evidence-based geriatric nursing protocols for best practice.

Kresevic DM. Assessment of function. In: Evidence-based geriatric nursing protocols for best practice. 3rd ed.; pp. 23-40. New York: Springer Publishing Company, 2008. [48 references].

1c.11 National Guideline Clearinghouse or other URL:
http://www.guideline.gov/summary/summary.aspx?doc_id=12259&nbr=006343&string=functional+AND+decline

1c.12 Rating of strength of recommendation (*also provide narrative description of the rating and by whom*):
The strength of the recommendation has not been rated.

1c.13 **Method for r**ating strength of recommendation (*If different from* USPSTF system*, also describe rating and how it relates to USPSTF*):

1c.14 Rationale for using this guideline over others:
These are the relevant guidelines registered with the National Guideline Clearinghouse that address preventing ADL decline.

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	1
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? **Rationale:**	1 Y ☐ N ☐

2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, <u>as specified</u>, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	Eval Rating

2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained? **S.2** If yes, provide web page URL: 2a. Precisely Specified	2a-specs C ☐ P ☐ M
2a.1 Numerator Statement (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome*):	

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable   7

Comment [k7]: USPSTF grading system http://www.ahrq.gov/clinic/uspstf/grades.htm: **A** - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B** - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C** - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D** - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I** - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

Comment [KP8]: 2a. The measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. The required data elements are of high quality as defined by NQF's Health Information Technology Expert Panel (HITEP) .

The numerator is the number of long-stay residents who have an MDS assessment (which may be an annual, quarterly, significant change, or significant correction) reporting a defined amount of decline when compared with a previous assessment (which may be an admission, annual, quarterly, significant change, or significant correction MDS 3.0 assessment). This would indicate an increase, when compared with a previous assessment, in the resident's need for help with a late-loss item as indicated by a higher score (coding convention is such that a higher score indicates the need for more help with a task). The need for increased assistance (suggesting decline in function) is identified if the score for at least one late-loss ADL item increases by two or more points or if the score for two or more of the late-loss ADLs items increase by one point; late-loss ADL items are bed mobility, transferring, eating, and toileting.

2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator***):**
The numerator data are from the target quarter MDS 3.0 assessment (which may be an annual, quarterly, significant change, or significant correction assessment) and refers to the ADL decline reported since a previous assessment (which may be an admission, annual, quarterly, significant change, or significant correction MDS 3.0 assessment).

2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions***):**
Residents are counted if they are long-stay residents, defined as residents whose length of stay is greater than 100 days. Residents who return to the nursing home following a hospital discharge will not have their stay reset to zero. Residents are counted if they reported having an increase in their need for help with late-loss ADLs. An increase is defined as an increase in two or more coding points in one late-loss ADL item or a one point increase in coding points in two or more late-loss ADL items. The comparison is made between the target quarter´s assessment (which may be an annual, quarterly or significant change or significant correction MDS 3.0 assessment) and the previous assessment (which may be an annual, quarterly or significant correction MDS 3.0 assessment). Higher score on an item indicates greater dependency. The ADL items for this measure are: 1. Bed mobility-G0110A1 2. Transferring-G0110B1 3. Eating-G0110H1 4. Toileting-G011011. Note. Values of 7 (occurred only once or twice) or 8 (did not occur) are recoded to be a value of 4.

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured***):**
The denominator includes all long-stay residents who received an annual, quarterly or significant change or correction MDS 3.0 assessment during the quarter and who did not meet the exclusion criteria.

2a.5 Target population gender: Male, Female
2a.6 Target population age range: Our intention in specifying all ages was to indicate that CMS does not intend to report the measures for age-specific segments of the nursing facility

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator***):**
Denominator data come from MDS 3.0 annual, quarterly or significant change or correction assessment conducted during each quarter (3-month period).

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):**
Residents are counted if they are long-stay residents defined as residents whose length of stay is greater than 100 days. Residents who return to the nursing home following a hospital discharge will not have their day count reset to zero. The target population includes all long-stay residents who had an annual, quarterly, significant change, significant correction, or discharge assessment during the selected quarter.

2a.9 Denominator Exclusions (*Brief text description of exclusions from the target population***):** These are the two types of assessments that might be completed upon admission. OBRA regulations require a full assessment within 14 days of admission. Medicare SNF payments require a Prospective Payment System (PPS) assessment. Newly admitted residents (identified by having either of these two types of admission assessments) are not included in the denominator as this represents their baseline status, not whether they have declined since admission.

Denominator exclusion criteria include the following:

Comment [k9]: 11 Risk factors that influence outcomes should not be specified as exclusions.
12 Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

• an OBRA admission assessment or a 5-day Prospective Payment System (PPS) target assessment,
• the resident is totally dependent in all four late-loss ADL items,
• the resident is comatose,
• the resident is receiving hospice care, or
• the resident does not meet the criteria for decline in late-loss ADLs (an increase by two or more points in one late-loss ADL, or increase of one point in two or more late-loss ADLs) based on the ADL data available, AND there is missing data on any of the four late-loss ADL items .

Long-stay facilities are excluded from public reporting if their sample includes fewer than 30 residents.

2a.10 Denominator Exclusion Details (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions*):
1. If the target MDS 3.0 assessment is an OBRA admission assessment (A0310.A = 01, indicating that the assessment is the admission assessment, conducted within 14 days of admission) or the assessment is a PPS assessment (A0310.B = 01, 02, 02, 03, 04, 05, 06 or 07) or there is missing data for any of the following (so that the measure cannot be accurately calculated)
2. All four late-loss ADL items indicate total dependence (all = 4, 7 or 8) (and so cannot decline further)
3. The resident is Comatose (B0100 = 1) (also expected to be totally dependent and unable to decline further)
4. Prognosis of life expectancy is less than 6 months (J1400=1 or missing)
5. Hospice care (O0100.K.2=1 or missing) (and so decline is anticipated)
6. the resident does not meet the criteria for decline in late-loss ADLs (an increase by two or more points in one late-loss ADL, or increase of one point in two or more late-loss ADLs) based on the ADL data available, AND there is missing data on any of the four late-loss ADL items.

2a.11 Stratification Details/Variables (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions*):
This is not applicable.

2a.12-13 Risk Adjustment Type: No risk adjustment necessary

2a.14 Risk Adjustment Methodology/Variables (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*):

2a.15-17 Detailed risk model available Web page URL or attachment:

2a.18-19 Type of Score: Ratio
2a.20 Interpretation of Score:
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps*):
For each facility, the number of residents meeting the numerator criteria and the number of residents meeting the denominator criteria are counted. The facility prevalence score is calculated as the number of residents in the facility during the selected quarter in the numerator divided by all residents during the selected quarter in the denominator.

2a.22 Describe the method for discriminating performance (*e.g., significance testing*):
Because the computed scores are not estimates, but include all residents who meet the measure criteria, in terms of discriminating performance, the computed scores can be used to make valid comparisons.

2a.23 Sampling (Survey) Methodology *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:
This is not applicable.

2a.24 Data Source (*Check the source(s) for which the measure is specified and tested*)
Electronic clinical data

2a.25 Data source/data collection instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*):
The data source or collection instrument is Nursing Home MDS 3.0.

2a.26-28 Data source/data collection instrument reference web page URL or attachment: URL
http://www.cms.hhs.gov/NursingHomeQualityInits/25_NHQIMDS30.asp#TopOfPage

2a.29-31 Data dictionary/code table web page URL or attachment: URL
http://www.cms.hhs.gov/NursingHomeQualityInits/25_NHQIMDS30.asp#TopOfPage,

2a.32-35 Level of Measurement/Analysis (*Check the level(s) for which the measure is specified and tested*)
Population: national, Facility/Agency

2a.36-37 Care Settings (*Check the setting(s) for which the measure is specified and tested*)
Nursing home (NH) /Skilled Nursing Facility (SNF)

2a.38-41 Clinical Services (*Healthcare services being measured, check all that apply*)

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample (*description of data/sample and size*): Two major tests were conducted of the reliability of the ADL decline measure; in addition, earlier analyses evaluated inter-rater reliability of the underlying MDS items in a more limited analysis. First, the MDS 2.0 measured items and the existing quality measure were tested in the Data Assessment and Verification (DAVE 2) project conducted by Abt Associates. This project used a nationwide sample of randomly selected nursing facilities using MDS assessments for the period from April 1 to December 31, 2006; 173 two-stage reviews were performed.(1)

Second, the University of Colorado used national facility-level quality measure data from third quarter (Q3) of 2003 through Q3 of 2006, which came from the Quality Improvement and Evaluation System (QIES) MDS Express Reports on the CMS Intranet; and Online Survey, Certification, and Reporting (OSCAR) data related to facility characteristics (e.g., state, resident census, number of beds, staffing) and certification survey results were downloaded from the QIES Workbench.(2) A 10% random sample of all Medicare-certified nursing facilities was also downloaded from MDS assessment records. Analyses were based on complete MDS data from January 2005 through March 2006, nearly complete data for April 2006, and partial data for May and June 2006.(2)

Earlier analyses used a sample of 219 facilities in six states and 5,758 residents to evaluate the inter-rater reliability on the individual ADL items used in the ADL decline quality measure. Researchers compared the ratings provided by research nurse ("gold standard") pairs and evaluated the percentage of agreement and provided kappa and weighted kappa statistics.(3)

1. Abt Associates, Inc.; Stepwise Systems, Inc.; Qualidigm. Data Assessment and Verification (DAVE 2) project—MDS two-stage discrepancy findings, April–December 2006. Cambridge, MA: Abt Associates, Inc, 2007.

2. Brega A, Hittle D, Goodrich G, Kramer A, Conway K, Levy C. Empirical review of publicly reported nursing home quality measures. Denver: Division of Health Care Policy and Research University of Colorado at Denver; Abt Associates, Inc, 2007.

3. Mor V, Angelelli J, Jones R, Roy J, Moore T, Morris J. Inter-rater reliability of nursing home quality indicators in the U.S. BMC Health Serv Res. 2003 November 4;3(1):20.

2b.2 Analytic Method (*type of reliability & rationale, method for testing*):
The DAVE 2 project used a two-stage cluster sample design to examine MDS reporting. A trained nurse reviewer selected a current resident with a recent assessment performed by the nursing facility within the past 14 days. In Stage 1 of this review, the nurse reviewer conducted a blind reassessment of the resident using standard MDS assessment and coding procedures (examination of the medical record, observation of the resident, interview of staff, resident, and family, and use of coding criteria). In Stage 2 of this assessment, the DAVE 2 nurse reviewer's assessment was compared to the corresponding nursing facility assessment and each discrepancy was reconciled, with the nursing home assessor and the nurse reviewer agreeing on the appropriate response. In addition to data entering the facility MDS code, the DAVE 2 code, and the reconciled code into the MDS-QC data entry software, the DAVE 2 nurse reviewer entered a "reason code" to attribute the cause of the discrepancy, per MDS item reviewed, to an established list of reasons.

2b
C ☐
P ☐
M
☐
N ☐

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

Comment [KP10]: 2b. Reliability testing demonstrates the measure results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.

Comment [k11]: 8 Examples of reliability testing include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing may address the data items or final measure score.

2b.3 Testing Results *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
As part of the DAVE 2 project, Abt Associates used item-level data to estimate reliability because measure-level reliability could not be assessed with data from a single time point. Hence, this is only a partial reliability test for the ADL decline measure. These items also have multiple response options, for which Abt determined that a 10% discrepancy rate would be acceptable. The two-stage review discrepancy rate for the response options for the individual late-loss ADL measures were substantially higher than 10%: ranging from 22.6% to 27.6%. However, the testing did not evaluate the extent of the discrepancy (i.e., how large the discrepancies were).

Evaluating the component items used in the current MDS 2.0 quality measure, there was high inter-rater reliability (comparing results obtained from research nurse pairs) on transferring, eating, toileting, and bed mobility (agreement was very high (96% to 98%, with weighted kappas of 0.85-0.91).

1. Abt Associates, Inc.; Stepwise Systems, Inc.; Qualidigm. Data Assessment and Verification (DAVE 2) project—MDS two-stage discrepancy findings, April-December 2006. Cambridge, MA: Abt Associates, Inc, 2007.

2. Mor V, Angelelli J, Jones R, Roy J, Moore T, Morris J. Inter-rater reliability of nursing home quality indicators in the U.S. BMC Health Serv Res. 2003 November 4;3(1):20.

2c. Validity testing

> **Comment [KP12]:** 2c. Validity testing demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.

2c.1 Data/sample *(description of data/sample and size)*: The data came from two sources: national facility-level quality measure data from Q3 of 2003 through Q3 of 2006, which came from the QIES MDS Express Reports on the CMS Intranet; and OSCAR data related to facility characteristics (e.g., state, resident census, number of beds, staffing) and certification survey results were downloaded from the QIES Workbench. A 10% random sample of all Medicare-certified nursing facilities was also downloaded from MDS assessment records. Analyses were based on complete MDS data from January 2005 through March 2006, nearly complete data for April 2006, and partial data for May and June 2006.

1. Brega A, Hittle D, Goodrich G, Kramer A, Conway K, Levy C. Empirical review of publicly reported nursing home quality measures. Denver: Division of Health Care Policy and Research University of Colorado at Denver; Abt Associates, Inc, 2007.

2. Brega A, Goodrich G, Nuccio E, Hittle D. Transition of publicly reported nursing home quality measures to MDS 3.0—draft. Denver: Division of Health Care Policy and Research University of Colorado at Denver, 2008.

2c.2 Analytic Method *(type of validity & rationale, method for testing)*:
The analysis evaluated measure validity in a number of ways to examine the expected positive influence of public reporting on quality of care, which is an assessment of the degree to which quality measure triggering rates have improved over time; evaluate convergent validity, which is an assessment of the correlation of the quality measure with all other measures; and determine if the quality measure triggering rate was influenced by factors that are unrelated to facility quality, which is an evaluation of seasonal variations in triggering rates across the 13 quarters of data. The analysis also computed descriptive statistics and conducted a one-way analysis of variance (ANOVA) for the measure to examine the amount of variance in triggering rates explained by the state where a facility was located.

> **Comment [k13]:** 9 Examples of validity testing include, but are not limited to: determining if measure scores adequately distinguish between providers known to have good or poor quality assessed by another valid method; correlation of measure scores with another valid indicator of quality for the specific topic; ability of measure scores to predict scores on some other related valid measure; content validity for multi-item scales/tests. Face validity is a subjective assessment by experts of whether the measure reflects the quality of care (e.g., whether the proportion of patients with BP < 140/90 is a marker of quality). If face validity is the only validity addressed, it is systematically assessed (e.g., ratings by relevant stakeholders) and the measure is judged to represent quality care for the specific topic and that the measure focus is the most important aspect of quality for the specific topic.

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
The ADL decline quality measure performed well on indicators of reportability and variability, but less well on stability. The measure was reportable for 81.8% of facilities and was associated with a reasonable degree of variability (see table below, which presents national data from the first quarter [Q1] of 2006).

See attached Table 1: Measure Variability Across Facilities.

The measure showed substantial instability in facility triggering rates over time, with 33.8% of facilities

2c
C ☐
P ☐
M ☐
N ☐

experiencing changes of three deciles or more in facility ranking from quarter to quarter. ADL decline was well correlated with one other quality measure (mobility decline, R = 0.61). In this study, the measure showed no evidence of improvement over time in this study (from Q3 of 2003 to Q1 of 2006). However, although the most recent national rate shows no improvement in the mean ADL decline nationally (15.7% of residents), the highest state mean is much lower (22% compared to 27% previously).(1) The measure also demonstrated substantial seasonal variation, suggesting that the measure is influenced by seasonal changes in resident case mix.

Results from the DAVE 2 project indicated that the four data elements on which the measure is based show substantial inter-rater discrepancies (responses differ between 23.1% to 27.6% of the time).(2) Although the DMINHo team attempted to develop a risk model that could account for such variation in resident characteristics, they were not successful in building a model that met their threshold for adequate predictive performance.

1. CMS. MDS quality measure/indicator report. 2008. Available from http://www.cms.hhs.gov/MDSPubQIandResRep/02_qmreport.asp#TopOfPage.

2. Abt Associates, Inc.; Stepwise Systems, Inc.; Qualidigm. Data Assessment and Verification (DAVE 2) project—MDS two-stage discrepancy findings, April–December 2006. Cambridge, MA: Abt Associates, Inc., 2007.

2d. Exclusions Justified

2d.1 Summary of Evidence supporting exclusion(s):
All long-stay residents for whom complete data exists are included, except for those who are fully dependent and hence cannot decline further. Hospice patients and those with a prognosis of 6 months or less are excluded because ADL decline is expected.(1)

2d.2 Citations for Evidence:
1. Chen JH, Chan D, Kiely DK, Morris JN, Mitchell SL. Terminal trajectories of functional decline in the long-term care setting. J Gerontol Med Sci. 2007;62A(5):M531-6.

2d.3 Data/sample (description of data/sample and size): This is not applicable.

2d.4 Analytic Method (type analysis & rationale):
This is not applicable.

2d.5 Testing Results (e.g., frequency, variability, sensitivity analyses):
This is not applicable.

2d
C ☐
P ☐
M
☐
N
☐
NA
☐

2e. Risk Adjustment for Outcomes/ Resource Use Measures

2e.1 Data/sample (description of data/sample and size): This is not applicable.

2e.2 Analytic Method (type of risk adjustment, analysis, & rationale):
This is not applicable.

2e.3 Testing Results (risk model performance metrics):
This is not applicable.

2e.4 If outcome or resource use measure is not risk adjusted, provide rationale: The data elements have not changed since the update to MDS 3.0 because many states were not prepared to make the requisite changes to their IT systems to convert to a new series of items for calculating the RUGS-based payments. There are only minor changes/clarifications to the instructions for completing the MDS 3.0 ADL items, and the addition of a rating option for "Activity occurred only once or twice" which will be recoded to "Activity did not occur" for QM calculation. Otherwise, the ADL item remained the same as in MDS 2.0.

The measure is not risk adjusted. Results from the empirical testing of risk adjustment models using the MDS 2.0 specifications were poor (i.e., R-square = 0.0054 [i.e., less than 0.1—the standard established in the

2e
C ☐
P ☐
M
☐
N
☐
NA
☐

Comment [KP14]: 2d. Clinically necessary measure exclusions are identified and must be:
•supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;
AND
•a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus;
 AND
•precisely defined and specified:
–if there is substantial variability in exclusions across providers, the measure is specified so that exclusions are computable and the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);
if patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that it strongly impacts performance on the measure and the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

Comment [k15]: 10 Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, sensitivity analyses with and without the exclusion, and variability of exclusions across providers.

Comment [KP16]: 2e. For outcome measures and other measures (e.g., resource use) when indicated:
•an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care;Error! Bookmark not defined. OR rationale/data support no risk adjustment.

Comment [k17]: 13 Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than adjusting out differences.

analyses]).(1) Analyses using risk adjustment in combination with multilevel modeling and empirical Bayes estimates had little impact on the resulting ADL decline rates, or the number of facilities flagged at the 90th percentile for this measure.(2) Using slightly different measures, focusing on an admission cohort and imputing values for the residents who left the facility prior to the first quarterly assessment, researchers analyzed various approaches to predicting ADL decline and found the R-squares for models restricted to individual characteristics were low, ranging from .04 to .12.(3)

1. Brega A, Hittle D, Goodrich G, Kramer A, Conway K, Levy C. Empirical review of publicly reported nursing home quality measures. Denver: Division of Health Care Policy and Research University of Colorado at Denver; Abt Associates, Inc, 2007.
2. Arling G, Lewis T, Kane RL, Mueller C, Flood S. Improving quality assessment through multilevel modeling: the case of Nursing Home Compare. Health Services Research. 2007 June;42:(3 Part 1):1177-99.
3. Phillips C, Shen S, Chen M, Sherman M. Evaluating nursing home performance indicators: An illustration exploring the impact of facilities on ADL change. Gerontologist. 2007;47(5);683-9.

2f. Identification of Meaningful Differences in Performance

2f.1 Data/sample from Testing or Current Use *(description of data/sample and size)*: The data came from two sources: national facility-level quality measure data from Q3 of 2003 through Q3 of 2006, which came from the QIES MDS Express Reports on the CMS Intranet; and OSCAR data related to facility characteristics (e.g., state, resident census, number of beds, staffing) and certification survey results were downloaded from the QIES Workbench. A 10% random sample of all Medicare-certified nursing facilities was also downloaded from MDS assessment records. Analyses were based on complete MDS data from January 2005 through March 2006, nearly complete data for April 2006, and partial data for May and June 2006.

2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance *(type of analysis & rationale)*:
Because the computed scores are not estimates, but include all residents who meet the measure criteria, in terms of discriminating performance, the computed scores can be used to make valid comparisons.

2f.3 Provide Measure Scores from Testing or Current Use *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*:
 An analytical team at the University of Colorado's Health Sciences Center examined the triggering rates for the current measure at the facility level. Below are the measure scores from testing or current use (description of scores [e.g., distribution by quartile, mean, median, standard deviation], identification of statistically significant and meaningfully differences in performance). The measure was reportable for 81.8% of facilities and was associated with a reasonable degree of variability. See attached Table 1: Measure Variability Across Facilities, which presents national data from Q1 of 2006.

The desired outcome cannot be determined because the rate of unavoidable decline associated with disease progression has not been established and may vary by case mix.

2f	
C	☐
P	☐
M	☐
N	☐

2g. Comparability of Multiple Data Sources/Methods

2g.1 Data/sample *(description of data/sample and size)*: This is not applicable.

2g.2 Analytic Method *(type of analysis & rationale)*:
This is not applicable.

2g.3 Testing Results *(e.g., correlation statistics, comparison of rankings)*:
This is not applicable.

2g	
C	☐
P	☐
M	☐
N	☐
NA	☐

2h. Disparities in Care

2h.1 If measure is stratified, provide stratified results *(scores by stratified categories/cohorts)*: This measure is not stratified.

2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:

2h	
C	☐
P	☐
M	☐
N	☐

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

---

Comment [KP18]: 2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.

Comment [k19]: 14 With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74% v. 75%) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall poor performance may not demonstrate much variability across providers.

Comment [KP20]: 2g. If multiple data sources/methods are allowed, there is demonstration they produce comparable results.

Comment [KP21]: 2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender);OR rationale/data justifies why stratification is not necessary or not feasible.

Although MDS 3.0 collects data on the resident's race, there are no current plans to stratify the measure by race because facilities tend to be homogenous by race, making disparities generally evident in the rating of the facility.(1, 2, 3) We plan to evaluate whether there are adequate numbers of individuals with or without significant cognitive impairment and with significantly different trajectories to stratify the sample accordingly. 1. Smith D, Feng Z, Zinn J, Mor V. 2008. Racial disparities in access to long-term care: the illusive pursuit of equity. J Health Polit Policy Law. 2008;33(5):861-81. 2. Smith D, Feng Z, Fennell M, Zinn J, Mor V. Separate and unequal: racial segregation and disparities in quality across U.S. nursing homes. Health Aff (Millwood). 2007;26(5):1448-1558. 3. Mor V, Berg K, Angelelli J, Gifford D, Morris J, Moore T. 2003. The quality of quality measurement in U.S. nursing homes. Geronotologist. 2003;43(Special Issue II):37-46.	**NA** ☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties*?	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met? **Rationale:**	**2** **C** ☐ **P** ☐ **M** ☐ **N** ☐

3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	Eval Rating

3a. Meaningful, Understandable, and Useful Information **3a.1 Current Use:** In use **3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*): Nursing Home Compare http://www.medicare.gov/NHCompare/Include/DataSection/Questions/SearchCriteriaNEW.asp?version=default&browser=IE%7C6%7CWinXP&language=English&defaultstatus=0&pagelist=Home&CookiesEnabledStatus=True **3a.3 If used in other programs/initiatives** (*If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years*): CMS expects that the quality measure will be used by nursing homes as a tool to evaluate their performance and develop quality improvement activities to prevent or minimize ADL decline. **Testing of Interpretability** (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*) **3a.4 Data/sample** (*description of data/sample and size*): A recent study examined whether consumers could accurately interpret the quality information given for all the measures reported by Nursing Home Compare.(1) Data were collected from 4,754 family members of nursing home residents. 1. Castle N. The Nursing Home Compare report card: consumers' use and understanding. J Aging Soc Policy.	**3a** **C** ☐ **P** ☐ **M** ☐ **N** ☐

Comment [KP22]: 3a. Demonstration that information produced by the measure is meaningful, understandable, and useful to the intended audience(s) for <u>both</u> public reporting (e.g., focus group, cognitive testing) <u>and</u> informing quality improvement (e.g., quality improvement initiatives). An important outcome that may not have an identified improvement strategy still can be useful for informing quality improvement by identifying the need for and stimulating new approaches to improvement.

2009;21(2):187-208.

3a.5 Methods *(e.g., focus group, survey, QI project)*:
A comprehension index was used to examine whether the information contained in Nursing Home Compare for each quality measure was understood by family members.

3a.6 Results *(qualitative and/or quantitative results and conclusions)*:
The study found that 31% of the consumers used the Internet to help them choose a nursing facility, 12% recalled using Nursing Home Compare, and, in general, the consumers' comprehension index scores were high, indicating a good understanding. The comprehension index for the ADL decline measure was 5.65 on a scale of 1 to 8, somewhat higher than the mean for all non-risk adjusted measures of 5.35.

3b/3c. Relation to other NQF-endorsed measures

3b.1 NQF # and Title of similar or related measures:
The proposed measure is intended to replace NQF #0182—Residents whose need for help with activities of daily living has increased, which is based on the MDS 2.0. The MDS 2.0 is being replaced by the MDS 3.0. Other related measures are: NQF #0430—Change in daily activity function as measured by the AM-PAC (Home Health) and NQF # 0175—Improvement in bed transferring (Home Health)

(for NQF staff use) Notes on similar/related endorsed or submitted measures:

3b. Harmonization
If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population):
3b.2 Are the measure specifications harmonized? If not, why?
The specifications for calculating the proposed measure have not changed from those used to calculate NQF #0182—Residents whose need for help with activities of daily living has increased (which is based on the MDS 2.0) except to reflect the item numbering in the MDS 3.0.

3c. Distinctive or Additive Value
3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:

5.1 If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability*?

Steering Committee: Overall, to what extent was the criterion, *Usability*, met?
Rationale:

3b
C ☐
P ☐
M ☐
N ☐
NA ☐

3c
C ☐
P ☐
M ☐
N ☐
NA ☐

3

3
C ☐
P ☐
M ☐
N ☐

Comment [KP23]: 3b. The measure specifications are harmonized with other measures, and are applicable to multiple levels and settings.

Comment [k24]: 16 Measure harmonization refers to the standardization of specifications for similar measures on the same topic (e.g., *influenza immunization* of patients in hospitals or nursing homes), or related measures for the same target population (e.g., eye exam and HbA1c for *patients with diabetes*), or definitions applicable to many measures (e.g., age designation for children) so that they are uniform or compatible, unless differences are dictated by the evidence. The dimensions of harmonization can include numerator, denominator, exclusions, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources.

Comment [KP25]: 3c. Review of existing endorsed measures and measure sets demonstrates that the measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare, is a more valid or efficient way to measure).

4. FEASIBILITY

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)

4a. Data Generated as a Byproduct of Care Processes

4a.1-2 How are the data elements that are needed to compute measure scores generated?

Eval Rating

4a
C ☐
P ☐

Comment [KP26]: 4a. For clinical measures, required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery. (e.g., BP recorded in the electronic record, not abstracted from the record later by other personnel; patient self-assessment tools, e.g., depression scale; lab values, meds, etc.)

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition), Coding/abstraction performed by someone other than person obtaining original information (E.g., DRG, ICD-9 codes on claims, chart abstraction for quality measure or registry)	M ☐ N ☐

4b. Electronic Sources

4b.1 Are all the data elements available electronically? (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*)
No

4b.2 If not, specify the near-term path to achieve electronic capture by most providers.

4b
C ☐
P ☐
M ☐
N ☐

4c. Exclusions

4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?
No

4c.2 If yes, provide justification.

4c
C ☐
P ☐
M ☐
N ☐
NA ☐

4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences

4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.
This measure has some weaknesses which could lead to inaccuracies in ratings reflecting case mix rather than care received.

Participants in the clinical review of the DMNHO project reported facility administrators and staff consider the ADL decline measure of limited value because it is not possible to distinguish decline, resulting from inadequate care from unavoidable decline.(1) State Medicaid payment policies may also affect reporting: more ADL decline was reported among facilities in states using MDS data for payment than in states that do not, although the researchers suggest this may result from an increased incentive to accurately report decline.(2)

Studies consistently show that ADL decline is associated with cognitive function (e.g., 3, 4) and other individual characteristics, suggesting a need for risk adjustment, which have been unsuccessful as of yet. The extent to which quality may explain variation in reported rates of ADL decline has not been established. A study of ADL decline in variously defined admission cohorts found that facility identity alone explained only 8% to 14% of the variation in ADL functional change; even when resident characteristics were added to facility identity, the total explained variation never exceeded 20%.(4) CMS studies to validate quality indicators found that facility identity explained 26% of change in ADL.(5)

However, there is consensus that it is important to evaluate ADL decline in long-stay nursing facility residents as maintaining function is a key goal of nursing facility care. This is reflected in the inclusion of the current MDS 2.0-based quality measure as part of the CMS 5 Star rating system.

1. Brega A, Hittle D, Goodrich G, Kramer A, Conway K, Levy C. Empirical review of publicly reported nursing home quality measures. Denver: Division of Health Care Policy and Research University of Colorado at Denver; Abt Associates, Inc, 2007.

2. Bellows NM, Halpin HA. MDS-based state Medicaid reimbursement and the ADL-decline quality indicator. Gerontologist. 2008;48(3):324-29.

3. Ang YH, Su SYL, Yap LKP, Ee CH. Functional decline of the elderly in a nursing home. Singapore Med J. 2006;47(3):219-224.

4d
C ☐
P ☐
M ☐
N ☐

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

16

4. Phillips C, Shen S, Chen M, Sherman M. Evaluating nursing home performance indicators: An illustration exploring the impact of facilities on ADL change. Gerontologist. 2007;47(5);683-89. 5. Morris JN, Moore T, Jones R, Mor V, Angelilli J, Berg K, et al. Validation of long-term and post-acute care quality indicators. Final draft report, Version 2. Cambridge MA: Abt Associates, 2002. This was not audited.	

4e. Data Collection Strategy/Implementation **4e.1** Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: The data collection method is already in operational use, and no issues are anticipated with these areas. **4e.2** Costs to implement the measure (*costs of data collection, fees associated with proprietary measures*): Data are collected as part of an existing process with no additional cost. **4e.3** Evidence for costs: This is not applicable. **4e.4** Business case documentation: The proposed measure relies on data from the MDS 3.0. As there is no change in the data collection method for the MDS 3.0 as compared with its predecessor, the MDS 2.0, we do not anticipate any additional burden to nursing facilities. MDS 2.0, and soon to be MDS 3.0, data are collected as part of an existing, federally mandated process used for payment and quality monitoring purposes.	4e C ☐ P ☐ M ☐ N ☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility*?	4
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? Rationale:	4 C ☐ P ☐ M ☐ N ☐

Comment [KP30]: 4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, etc.) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).

RECOMMENDATION	
(for NQF staff use) Check if measure is untested and only eligible for time-limited endorsement.	Time-limited ☐
Steering Committee: Do you recommend for endorsement? Comments:	Y ☐ N A ☐

CONTACT INFORMATION	
Co.1 Measure Steward (Intellectual Property Owner) **Co.1** <u>Organization</u> Centers for Medicare & Medicaid Services, 7500 Security Boulevard, Mail Stop S3-02-01, Baltimore, Maryland, 21244-1850 **Co.2** <u>Point of Contact</u> Judith, Tobin, PT, MBA, Judith.Tobin@cms.hhs.gov, 410-786-6892-	
Measure Developer If different from Measure Steward **Co.3** <u>Organization</u> RTI International, 1440 Main Street, Suite 300, Waltham, Massachusetts, 02451-1623	

Co.4 Point of Contact
Roberta, Constantine, RN, MBA, PhD, rconstantine@rti.org, 781-434-1700-1711

Co.5 Submitter If different from Measure Steward POC
Roberta, Constantine, RN, MBA, PhD, rconstantine@rti.org, 781-434-1700-1711, RTI International

Co.6 Additional organizations that sponsored/participated in measure development

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
See attached Table 2: Nursing Home Quality Measures Technical Expert Panel (January 2009).
This technical expert panel met during 2 days in January of 2009 to review the environmental scan of the current quality measures and make recommendations regarding their transition from MDS 2.0 to MDS 30.

Ad.2 If adapted, provide name of original measure: This measure was adapted from the measure of the same name derived from MDS 2.0 data.
Ad.3-5 If adapted, provide original specifications URL or attachment
http://www.qualitynet.org/dcs/ContentServer?cid=1138050766910&pagename=Medqic%2FOtherResource%2FOtherResourcesTemplate&c=OtherResource

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 Year the measure was first released: 2002
Ad.7 Month and Year of most recent revision: 02, 2010
Ad.8 What is your frequency for review/update of this measure? Every 3 years
Ad.9 When is the next scheduled review/update for this measure? 02, 2013

Ad.10 Copyright statement/disclaimers:

Ad.11 -13 Additional Information web page URL or attachment: Attachment Activities of Daily Living tables_FINAL.doc

Date of Submission (*MM/DD/YY*): 07/12/2010

Project Name: NQF Nursing Home Project

Measure Title: Percent of Residents Whose Need for Help with Activities of Daily Living Has Increased (Long Stay)

Planned Date of Measure Submission: March 19, 2010

Steward Name:

Point of Contact
Judith C. Tobin, PT, MBA
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Mail Stop S3-02-01
Baltimore, MD 21244-1850
410-786-6892
Judith.Tobin@cms.hhs.gov

Developer/Submitter Name:

RTI International
Roberta Constantine, RN, MBA, PhD
1440 Main Street, Suite 310
Waltham, MA 02451-1623
781-434-1711
rconstantine@rti.org

Table 1. Measure Variability Across Facilities

Quality Measure (QM)	N of Facilities[1]	Mean	Std Dev	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile	Facilities with QM = 0%
ADL Decline	13,253	15.5%	8.5%	5.9%	9.5%	14.1%	20.2%	27.0%	0.9%

Table 2. Nursing Home Quality Measures Technical Expert Panel (January 2009)

Name	Title	Affiliation
Barbara Anglin, RN	Program Services Consultant	American Association of Nurse Assessment Coordinators (AANAC)
Bonnie Burak-Danielson, MSM, EXP, LPTA	Rehab Manager of Reimbursement	Spaulding Rehab Network
Sarah Burger, MPH, RN	Senior Advisor and Coordinator	Coalition of Geriatric Nursing Organizations The John A. Hartford Institute for Geriatric Nursing
Diane Carter, MSN, RN, CS	President	AANAC
Kate Dennison, RN, RAC-MT	Minimum Data Set (MDS) Coordinator	The Cedars
Mary Ellard, RN, MPA/H, RAC-CT	Clinical Assessment Specialist	Five Star Quality Care, Inc.
Sandy Fitzler, RN	Senior Director of Clinical Services	American Health Care Association
David F. Hittle, PhD	Assistant Professor	Division of Health Care Policy and Research University of Colorado Denver, School of Medicine
Steve Levenson, MD, CMD	Multi-Facility Medical Director, Baltimore, MD	
Carol Maher, RN-BC, RAC-CT	Director of Clinical Reimbursement	Ensign Facilities Services
Barbara Manard, PhD	Vice President, Long Term Care/Health Strategies	American Association of Homes and Services for the Aging
Debra Saliba, MD, MPH	Anna and Harry Borun Chair in Geriatrics and Gerontology at UCLA Research Physician VA GLAHS GRECC Director of UCLA/JHA Borun Center for Gerentological Research Senior Natural Scientist RAND Health	University of California, Los Angeles (UCLA), Veterans Affairs (VA), RAND Corporation
Eric Tangalos, MD	Professor of Medicine	Mayo Clinic

Jacqueline Vance, RNC, CDONA/LTC	Director of Clinical Affairs	(American Medical Directors Association) AMDA
Mary Van de Kamp, MS/CCC-SLP	Vice President, Clinical Rehabilitation	People*first* Rehabilitation
Charlene Harrington, PhD, RN, FAAN*	Professor Emeritus	University of California, San Francisco Fellow in the American Academy of Nursing

Measure #/Title/Steward	
NH-022-10: Percent of Residents Whose Need for Help with Activities of Daily Living Has Increased (Long Stay) (Centers for Medicare & Medicaid Services)	
Description: This measure is based on data from the MDS 3.0 assessment of long-stay nursing facility residents and reports the percentage of all long-stay residents in a nursing facility whose need for help with late-loss Activities of Daily Living (ADLs), as reported in the target quarter's assessment, increased when compared with a previous assessment. The four late-loss ADLs are: bed mobility, transferring, eating, and toileting. This measure is calculated by comparing the change in each item between the target MDS assessment (which may be an annual, quarterly or significant change or correction assessment) and a previous assessment (which may be an admission, annual, quarterly or significant change or correction assessment).	
Initial In-Person Vote: Recommended for endorsement with conditions – 19 Not present - 1	
Steering Committee Questions/Conditions for Measure Developer:	**Response from Measure Developer**
The definition of long-stay residents needs to be clarified	Long-stay residents are defined as those greater than 100 days. Residents who return to the nursing home following a hospital discharge will not have their stay reset to zero.