September 2012

Nursing Home MDS 3.0 Quality Measures: Final Analytic Report

Deliverables 21, 25-27

Prepared for

Cheryl Wiseman, Capt, USPHS, MPH, MS
Centers for Medicare & Medicaid Services
Division of Ambulatory and Post Acute Care
Mail Stop C2-21-16
7500 Security Boulevard
Baltimore, MD 21244-1850

Prepared by

Laura Smith, PhD
Nan Tracy Zheng, PhD
Karen Reilly, ScD
Stephanie Kissam, MPH
Franziska Rokoske, PT, MS
Daniel Barch, MS
Yevgeniya Kaganova, PhD
Audrey Etlinger, MPP
Joshua Manning, BS

RTI International
1440 Main Street, Suite 310
Waltham, MA 02451-1623

RTI Project Number 0211942.100.300

RTI International is a trade name of Research Triangle Institute.

This page intentionally left blank.

Nursing Home MDS 3.0 Quality Measures: Final Analytic Report

by
Laura Smith, PhD
Nan Tracy Zheng, PhD
Karen Reilly, ScD
Stephanie Kissam, MPH
Franziska Rokoske, PT, MS
Daniel Barch, MS
Yevgeniya Kaganova, PhD
Audrey Etlinger, MPP
Joshua Manning, BS

Scientific Reviewers: Joshua M. Wiener, PhD; Zhanlian Feng, PhD

Federal Project Officer: Cheryl Wiseman, Capt USPHS, MPH, MS

RTI International

CMS Contract No. HHSM-500-2008-00021I

September 2012

This project was funded by the Centers for Medicare & Medicaid Services under contract no. HHSM-500-2008-00021I. The statements contained in this report are solely those of the authors and do not necessarily reflect the views or policies of the Centers for Medicare & Medicaid Services. RTI assumes responsibility for the accuracy and completeness of the information contained in this report.

RTI International is a trade name of Research Triangle Institute.

This page intentionally left blank.

CONTENTS

EXECUTIVE SUMMARY 1
E.1 Introduction 1
E.2 Data, Analytic Approach and Short and Long Population 1
Data 1
Analytic Approach 1
Short-Stay and Long-Stay Population 3
E.3 Key Findings from Testing 4
Variability 4
Reportability 5
Reliability 6
Validity 6
Risk-Adjustment Analyses 8
Quality Measure Specific Analyses 9
E.4 Summary 22

SECTION 1 INTRODUCTION 25
1.1 Statement of Purpose 25
1.2 History of MDS 3.0 QM Development 25
1.3 Analytic Approach and Report Organization 30
1.3.1 Overview 30
1.3.2 Data 31
1.3.3 Analytic Approach 31
1.3.4 Report Organization 34

SECTION 2 DESCRIPTION OF DATA SET AND SHORT-STAY/LONG-STAY RESIDENT POPULATIONS 35
2.1 Introduction of Long-Stay and Short-Stay Definitions 35
2.2 History of Long-Stay and Short-Stay Definitions 35
2.3 Detailed Definitions of Long Stay and Short Stay 36
2.3.1 Short-Stay Definition 36
2.3.2 Long-Stay Definition 37
2.3.3 Other Key Definitions 37
2.4 Analyses of Short-Stay and Long-Stay Resident Populations 38
2.4.1 Descriptive Statistics 38
2.4.2 Relative Proportions of Short-Stay and Long-Stay Residents within Facilities 39
2.4.3 Distribution of Length of Stay and Resident Characteristics by Long Stay and Short Stay 40

2.5 Short-Stay Population Analyses 43
2.5.1 Distribution of Length of Stay and Resident Characteristics—Short-Stay Analyses 44
2.5.2 Short-Stay Residents Who Die Before Accumulating 100 CDIF 48
2.5.3 Early Long-Stay Residents 49
2.6 Summary and Conclusions 51

SECTION 3 ANALYSES OF 16 NQF-ENDORSED MEASURES 52
3.1 Variability 52
3.2 Reportability for Each QM and Impact of Discharge Assessment on Reportability 56
3.3 Reliability 58
3.4 Validity Analysis—Impact of Discharge Assessment on QM Scores 62
3.5 Validity Analysis—Variation by State 65
3.6 Summary of Risk-Adjustment Strategies 68
3.7 Summary 72

SECTION 4 QUALITY MEASURE–SPECIFIC ANALYSES 74
4.1 QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) 74
4.1.1 Summary of Findings 74
4.1.2 Background and Introduction to QM 75
4.1.3 Descriptive Statistics 77
4.1.4 Findings from Testing 80
4.1.5 Risk Adjustment 89
4.2 QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) 90
4.2.1 Summary of Findings 90
4.2.2 Background and Introduction to QM 91
4.2.3 Descriptive Statistics 94
4.2.4 Findings from Testing 98
4.2.5 Risk Adjustment 108
4.3 QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay) 114
4.3.1 Summary of Findings 114
4.3.2 Background and Introduction to QM 114
4.3.3 Descriptive Statistics 117
4.3.4 Findings from Testing 118
4.3.5 Risk Adjustment 128
4.4 QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) 128

4.4.1 Summary of Findings....128
4.4.2 Background and Introduction to QM....129
4.4.3 Descriptive Statistics....132
4.4.4 Findings from Testing....133
4.4.5 Risk Adjustment....141
4.5 QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)....141
4.5.1 Summary of Findings....141
4.5.2 Background and Introduction to QM....142
4.5.3 Descriptive Statistics....144
4.5.4 Findings from Testing....146
4.5.5 Risk Adjustment....153
4.6 QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)....153
4.6.1 Summary of Findings....153
4.6.2 Background and Introduction to QM....154
4.6.3 Descriptive Statistics....157
4.6.4 Findings from Testing....160
4.6.5 Risk Adjustment....168
4.7 QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)....171
4.7.1 Summary of Findings....171
4.7.2 Background and Introduction to QM....172
4.7.3 Descriptive Statistics....175
4.7.4 Findings from Testing....178
4.7.5 Risk Adjustment....186
4.8 QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)....187
4.8.1 Summary of Findings....187
4.8.2 Background and Introduction to QM....188
4.8.3 Descriptive Statistics....190
4.8.4 Findings from Testing....192
4.8.5 Risk Adjustment....201
4.9 QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)....201
4.9.1 Summary of Findings....201
4.9.2 Background and Introduction to QM....201
4.9.3 Descriptive Statistics....204
4.9.4 Findings from Testing....205

4.9.5 Risk Adjustment....213
4.10 QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay)....214
4.10.1 Summary of Findings....214
4.10.2 Background and Introduction to QM....214
4.10.3 Descriptive Statistics....217
4.10.4 Findings from Testing....218
4.10.5 Risk Adjustment....226
4.11 QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)....226
4.11.1 Summary of Findings....226
4.11.2 Background and Introduction to QM....227
4.11.3 Descriptive Statistics....230
4.11.4 Findings from Testing....232
4.12.5 Risk Adjustment....240
4.12 QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)....241
4.12.1 Summary of Findings....241
4.12.2 Background and Introduction to QM....242
4.12.3 Descriptive Statistics....244
4.12.4 Findings from Testing....246
4.12.5 Risk Adjustment....254
4.13 QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)....257
4.13.1 Summary of Findings....257
4.13.2 Background and Introduction to QM....257
4.13.3 Descriptive Statistics....259
4.13.4 Findings from Testing....262
4.13.5 Risk Adjustment....269
4.14 QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)....269
4.14.1 Summary of Findings....269
4.14.2 Background and Introduction to QM....270
4.14.3 Descriptive Statistics....272
4.14.4 Findings from Testing....276
4.14.5 Risk Adjustment....283
4.15 QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)....284
4.15.1 Summary of Findings....284
4.15.2 Background and Introduction to QM....284

4.15.3 Descriptive Statistics....288
4.15.4 Findings from Testing....289
4.15.5 Risk Adjustment....297
4.16 QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) 297
4.16.1 Summary of Findings....297
4.16.2 Background and Introduction to QM....297
4.16.3 Descriptive Statistics....301
4.16.4 Findings from Testing....302
4.16.5 Risk Adjustment....314

SECTION 5 CONCLUSIONS....315
5.1 Summary....318

REFERENCES....319

List of Figures

Figure 2-1 Facility-level proportion of short-stay residents by facility volume....40
Figure 2-2 Discrete probability density of cumulative number of days in facility: short-stay resident episodes....45
Figure 4.1-1 Seasonal (quarterly) variation in QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)....89
Figure 4.2-2 Seasonal (quarterly) variation in QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)....108
Figure 4.3-1 Seasonal (quarterly) variation in QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)....128
Figure 4.4-1 Seasonal (quarterly) variation in QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)....141
Figure 4.5-1 Seasonal (quarterly) variation in QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)....153
Figure 4.6-1 Seasonal (quarterly) variation in QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)....168
Figure 4.7-1 Seasonal (quarterly) variation in QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)....186
Figure 4.8-1 Seasonal (quarterly) variation in QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay)....200
Figure 4.9-1 Seasonal (quarterly) variation in QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)....213
Figure 4.10-1 Seasonal (quarterly) variation in QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay)....226
Figure 4.11-1 Seasonal (quarterly) variation in QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)....240

Figure 4.12-1 Seasonal (quarterly) variation in QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 253
Figure 4.13-1 Seasonal (quarterly) variation in QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) 269
Figure 4.14-1 Pattern comparison before and after implementing MDS 3.0 for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) 283
Figure 4.15-1 Seasonal (quarterly) variation in QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) 296
Figure 4.16-1 Seasonal (quarterly) variation in QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) 314

List of Tables

Table 1-1 CMS NQF endorsed quality measures: names and key characteristics 28
Table 1-2 NQF evaluation criteria for reliability and validity* 30
Table 2-2 Proportion of target assessments by type of assessment variables by short-stay and long-stay definition by record type 39
Table 2-3 Facility-level proportion of short-stay residents, fourth quarter of 2011 40
Table 2-4 Length-of-stay (cumulative number of days in facility) distributions for short-stay and long-stay resident episodes, fourth quarter of 2011 41
Table 2-5 Resident characteristics by short-stay and long-stay 42
Table 2-6 Short-stay resident characteristics by stratified by type of the initial assessment.....46
Table 2-7 Distribution of resident's overall expectation of care by short and long stay 48
Table 2-8 Facility distribution of percent of short-stay residents who die before accumulating 100 cumulative days in facility 49
Table 2-9 Impact of excluding early long-stay residents on short-stay quality measure scores, time period 50
Table 3-1 Quality measure score distribution by QM 54
Table 3-2 Number and percent of nursing homes able to report quality measures with and without discharge assessments included in quality measure calculations 56
Table 3-3 Reliability: proportions of facilities with declining or improving quality measure scores from Quarter 3 to Quarter 4, 2011 59
Table 3-4 Facility-level quality measure score changes after excluding discharge assessments 63
Table 3-5 Proportion of variance attributed to state-by-state differences for quality measures.....66
Table 3-6 Summary of MDS 3.0 quality measures and risk-adjustment techniques used (short stay) 69
Table 3-7 Summary of MDS 3.0 quality measures and risk-adjustment techniques used (long stay) 69
Table 3-8 Summary table of quality measures with denominator exclusion/sample restriction used and rationale (short stay) 70
Table 3-9 Summary table of quality measures with denominator exclusion/sample restriction used and rationale (long stay) 71

Table 4.1-1 Target assessments used in the calculation of QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) 78
Table 4.1-2 Frequency and percentage of responses on items used to define the numerator for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) 79
Table 4.1-3 QM score distribution for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) 80
Table 4.1-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) 82
Table 4.1-5 Distribution of change in facility rank from one quarter to the next, QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) 82
Table 4.1-6 Change in facility score from one quarter to the next, QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) 84
Table 4.1-7 Missing data compared across selected resident characteristics 86
Table 4.1-8 Distribution of facility-level missing rate for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) 87
Table 4.1-9 Distribution of facility-level missing rate for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) by quartile of QM score 87
Table 4.1-10 Distribution of facility-level score change after excluding discharge assessment for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) 88
Table 4.2-1 Target assessments used in the calculation of QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) 95
Table 4.2-2 Frequencies of new and worsened pressure ulcers since prior assessment as reported on target assessments in QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay), by stage, covariate exclusions not applied 97
Table 4.2-3 Frequencies by stage of pressure ulcers, by whether target assessment is in QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) numerator, by stage, covariate exclusions not applied 99
Table 4.2-4 Frequencies of pressure ulcers that were present on admission, by whether the target assessment is in QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) numerator, by stage, covariate exclusions not applied 100
Table 4.2-5 QM score distribution for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) 101
Table 4.2-6 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) 101
Table 4.2-7 Distribution of change in facility rank from one quarter to the next for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) 102

Table 4.2-8 Change in facility score from one quarter to the next for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay).....104
Table 4.2-9 Distribution of facility-level missing rate for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay).....106
Table 4.2-10 Distribution of facility-level missing rate for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) 107
Table 4.2-11 Distribution of facility-level score change after excluding discharge assessment for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) 107
Table 4.2-12 Frequencies on covariate values for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay).....109
Table 4.2-13 Impact of varying combination of covariates in risk adjustment for this measure frequency of covariate combinations used for risk adjustment: QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay).....110
Table 4.2-14 Intercepts and coefficients by quarter of data for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) 112
Table 4.2-15 Intercepts and coefficients for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)—Comparison between single-level model and hierarchical generalized linear model 113
Table 4.3-1 Target assessments used in the calculation of QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) 117
Table 4.3-2 Frequency and percentage of responses on items used to define the numerator for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay).....118
Table 4.3-3 Score distribution for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) 119
Table 4.3-3a Score distribution for submeasure QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay).....120
Table 4.3-3b Score distribution for submeasure to QM #0680B Percent of Residents Who Were Offered and Declined the Seasonal Influenza Vaccine.....120
Table 4.3-3c Score distribution for submeasure QM #0680C Percent of Residents Who Did Not Receive, due to Medical Contraindication, the Seasonal Influenza Vaccine 120
Table 4.3-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) 121
Table 4.3-5 Distribution of change in facility rank from one quarter to the next for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay).....123

Table 4.3-6 Change in facility score from one quarter to the next for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) 124
Table 4.3-7 Distribution of facility-level imputation rate for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) by quartile of QM score 126
Table 4.3-8 Distribution of facility-level score change after excluding discharge assessment for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) 127
Table 4.4-1 Target assessments used in the calculation of QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) 132
Table 4.4-2 Responses to O0300a (PPV status is up to date) and O0300b (reason for not receiving pneumonia vaccine) on target assessments 132
Table 4.4-3 QM score distribution for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) 134
Table 4.4-3a QM score distribution for submeasure QM #0682A Percent of Residents Who Received the Pneumococcal Vaccine (Short Stay) 134
Table 4.4-3b QM score distribution for submeasure QM #0682B Percent of Residents Who Were Offered and Declined the Pneumococcal Vaccine (Short Stay) 134
Table 4.4-3c QM score distribution for submeasure QM #0682C Percent of Residents Who Did Not Receive, due to Medical Contraindication, the Pneumococcal Vaccine (Short Stay) 135
Table 4.4-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) 136
Table 4.4-5 Distribution of change in facility rank from one quarter to the next for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) 137
Table 4.4-6 Changes in facility scores from one quarter to the next for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) 138
Table 4.4-7 Distribution of facility-level score change after excluding discharge assessment for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) 140
Table 4.5-1 Target assessments used in the calculation of QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) 145
Table 4.5-2 Percentage of resident episodes in which at least one assessment in the look-back period indicated a fall resulting in major injury 145
Table 4.5-3 QM score distribution for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)** 146
Table 4.5-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) 147

Table 4.5-5 Distribution of change in facility rank from one quarter to the next, QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) 148
Table 4.5-6 Change in facility score from one quarter to the next, QM #0674 Percent of Residents Experiencing One of More Falls With Major Injury (Long Stay) 149
Table 4.5-7 Distribution of facility-level missing rate for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) 151
Table 4.5-8 Distribution of facility-level missing rate for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) by Quartile of QM score 151
Table 4.5-9 Distribution of facility-level score change after excluding discharge assessment for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) 152
Table 4.6-1 Target assessments used in the calculation of QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 158
Table 4.6-2 Responses to items important to calculating numerator for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 159
Table 4.6-3 QM score distribution for QM #0677 Percentage of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 160
Table 4.6-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 161
Table 4.6-5 Distribution of change in facility rank from one quarter to the next, QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 162
Table 4.6-6 Change in facility score from one quarter to the next, QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 163
Table 4.6-7 Missing data compared across selected resident characteristics 166
Table 4.6-8 Distribution of facility-level missing rate for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 166
Table 4.6-9 Distribution of facility-level missing rate for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) by quartile of QM score 167
Table 4.6-10 Distribution of facility-level score change after excluding discharge assessment for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 167
Table 4.6-11 Frequency of covariate used for risk adjustment: QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 169
Table 4.6-12 Intercepts and coefficients by quarter of data for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) 170
Table 4.6-13 Intercepts and coefficients by quarter of data for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)—Comparison between single-level model and hierarchical generalized linear model 171
Table 4.7-1 Target assessments used in the calculation of QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) 176

Table 4.7-2 Pressure ulcer frequency on target assessments by stage of pressure ulcer 177
Table 4.7-3 Pressure-ulcer profiles for residents included in the numerator of QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay).....178
Table 4.7-4 QM score distribution for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) 179
Table 4.7-5 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay).....180
Table 4.7-6 Distribution of change in facility rank from one quarter to the next, QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay).....181
Table 4.7-7 Change in facility score from one quarter to the next, QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) 182
Table 4.7-8 Distribution of facility-level missing rate for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) 184
Table 4.7-9 Distribution of facility-level missing rate for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) by quartile by QM score 185
Table 4.7-10 Distribution of facility-level score change after excluding discharge assessment for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay).....185
Table 4.8-1 Target assessments used in the calculation of QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) 190
Table 4.8-2 Frequency and percentage responses on items used to define the numerator for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay).....191
Table 4.8-3 QM score distribution for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay).....192
Table 4.8-3a QM score distribution for submeasure QM #0681A Percent of Residents Who Received the Seasonal Influenza Vaccine (Long Stay) 192
Table 4.8-3b QM score distribution for submeasure QM #0681B Percent of Residents Who Were Offered and Declined the Seasonal Influenza Vaccine (Long Stay) 193
Table 4.8-3c QM score distribution for submeasure QM #0681C Percent of Residents Who Did Not Receive, Due to Medical Contraindication, the Seasonal Influenza Vaccine (Long Stay) 193
Table 4.8-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay).....194
Table 4.8-5 Distribution of change in facility rank from one quarter to the next, QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) 195
Table 4.8-6 Change in facility score from one quarter to the next, QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay) 196
Table 4.8-7 Distribution of facility-level missing rate for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay) 198

Table 4.8-8 Distribution of facility-level missing rate for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay) by quartile of QM score 199
Table 4.8-9 Distribution of facility-level score change after excluding discharge assessment for QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) 199
Table 4.9-1 Target assessments used in the calculation of QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) 204
Table 4.9-2 Responses to O0300a (Pneumonia Vaccine Up-to-Date) and O0300b (Reason for Not Receiving Pneumonia Vaccine) on target assessments....205
Table 4.9-3 QM score distribution for QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)....206
Table 4.9-3a QM score distribution for submeasure QM #0683A Percent of Residents Who Received the Pneumococcal Vaccine (Long Stay) 206
Table 4.9-3b QM score distribution for submeasure QM #0683B Percent of Residents Who Were Offered and Declined the Pneumococcal Vaccine (Long Stay)....206
Table 4.9-3c QM score distribution for submeasure QM #0683C Percent of Residents Who Did Not Receive, Due to Medical Contraindication, the Pneumococcal Vaccine (Long Stay) 207
Table 4.9-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0683 Percent of Long-stay Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)....208
Table 4.9-5 Distribution of change in facility rank from one quarter to the next, QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)....208
Table 4.9-6 Changes in facility scores from one quarter to the next, QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay) 210
Table 4.9-7 Distribution of facility-level score change after excluding discharge assessment for QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)....212
Table 4.10-1 Target assessments used in the calculation of QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay)....217
Table 4.10-2 Frequency and percentage responses on items used to define the numerator for QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay) 218
Table 4.10-3 QM score distribution for QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay) 219
Table 4.10-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay)....220
Table 4.10-5 Distribution of change in facility rank from one quarter to the next for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) 221
Table 4.10-6 Change in facility score from one quarter to the next for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) 222

Table 4.10-7 Distribution of facility-level missing rate for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) 224
Table 4.10-8 Distribution of facility-level missing rate for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) by quartile of QM score 224
Table 4.10-9 Distribution of facility-level score change after excluding discharge assessment for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) 225
Table 4.11-1 Target assessments used in the calculation of QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) 231
Table 4.11-2 Responses to H0300 (Bowel Incontinence) and H0400 (Bladder Incontinence) on target assessments not otherwise excluded from analysis 232
Table 4.11-3 QM score distribution for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) 233
Table 4.11-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) 234
Table 4.11-5 Distribution of change in facility rank from one quarter to the next for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) 235
Table 4.11-6 Change in facility score from one quarter to the next for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) 236
Table 4.11-7 Distribution of facility-level missing rate for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) 238
Table 4.11-8 Distribution of facility-level missing rate for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) by quartile of QM score 239
Table 4.11-9 Distribution of facility-level score change after excluding discharge assessment for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) 239
Table 4.12-1 Target assessments used in the calculation of QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 245
Table 4.12-2 Responses to H0100a (Have Indwelling Catheters) on target assessments not otherwise excluded from analysis 245
Table 4.12-3 QM score distribution for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 246
Table 4.12-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 247
Table 4.12-5 Distribution of change in facility rank from one quarter to the next for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 248

Table 4.12-6 Change in facility score from one quarter to the next for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 249
Table 4.12-7 Distribution of facility-level missing rate for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 251
Table 4.12-8 Distribution of facility-level missing rate for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) by quartile of QM score 252
Table 4.12-9 Distribution of facility-level score change after excluding discharge assessment for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 253
Table 4.12-10 Frequency and impact of covariates used for risk adjustment: QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 254
Table 4.12-11 Intercepts and coefficients by quarter of data for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) 255
Table 4.12-12 Intercepts and coefficients in Quarter 4, 2011 for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay): comparison between single-level model and hierarchical generalized linear model 256
Table 4.13-1 Target assessments used in the calculation of QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) 260
Table 4.13-2 Frequency of items used to calculate QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) 261
Table 4.13-3 QM score distribution for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) 262
Table 4.13-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) 263
Table 4.13-5 Distribution of change in facility rank from one quarter to the next, QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) 264
Table 4.13-6 Change in facility scores from one quarter to the next, QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) 265
Table 4.13-7 Distribution of facility-level missing rate for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) 267
Table 4.13-8 Distribution of facility-level missing rate for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) by quartile of QM score 267
Table 4.13-9 Distribution of facility-level score change after excluding discharge assessment for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) 268
Table 4.14-1 Target assessments used in the calculation of QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) 273

Table 4.14-2 Frequency and percentage of responses on items important for the numerator of QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) 274
Table 4.14-3 QM score distribution for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)....276
Table 4.14-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Have Increased (Long Stay) 277
Table 4.14-5 Distribution of change in facility rank from one quarter to the next, QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) 278
Table 4.14-6 Change in facility score from one quarter to the next, QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) 279
Table 4.14-7 Distribution of facility-level missing rate for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) 281
Table 4.14-8 Distribution of facility-level missing rate for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) by quartile of QM score....282
Table 4.14-9 Distribution of facility-level score change after excluding discharge assessment for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) 282
Table 4.15-1 Target assessments used in the calculation of QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)....288
Table 4.15-2 Responses to K0300, Loss of 5% or more in the last month or loss of 10% or more in last 6 months....289
Table 4.15-3 QM score distribution for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)* 290
Table 4.15-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)....291
Table 4.15-5 Distribution of change in facility rank from one quarter to the next, QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) 292
Table 4.15-6 Change in facility scores from one quarter to the next, QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)....293
Table 4.15-7 Distribution of facility-level missing rate for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)....295
Table 4.15-8 Distribution of facility-level missing rate for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) by quartile of QM score 295
Table 4.15-9 Distribution of facility-level score change after excluding discharge assessment for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)....296
Table 4.16-1 Target assessments used in the calculation of QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)....301

Table 4.16-2 Frequency of items that count toward numerator of QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) 303
Table 4.16-3 QM score distribution for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) 306
Table 4.16-4 Change in reportability based on inclusion or exclusion of discharge assessments for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) 308
Table 4.16-5 Distribution of change in facility rank from one quarter to the next, QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) 309
Table 4.16-6 Change in facility scores from one quarter to the next, QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) 310
Table 4.16-7 Distribution of facility-level missing rate for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) 312
Table 4.16-8 Distribution of facility-level missing rate for QM #0690 Percent of Residents Who Have Depressive Symptoms by quartile of QM score (Long Stay) 312
Table 4.16-9 Distribution of facility-level score change after excluding discharge assessment for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) 313

EXECUTIVE SUMMARY

E.1 Introduction

The Omnibus Budget Reconciliation Act of 1987 (OBRA '87) mandated the development of a resident assessment instrument capable of measuring nursing home residents' health and quality of life. A result of this legislation was the implementation of a standardized assessment instrument, the Minimum Data Set (MDS), required for all residents and patients receiving care in Medicare- and Medicaid-certified nursing homes. The MDS was designed to facilitate achievement and maintenance of the resident's highest practical level of well-being by identifying and documenting individual resident's needs and strengths, and incorporating this information into the resident's care plan (Centers for Medicare & Medicaid Services, 2010b).

On October 1, 2010, the Centers for Medicare and Medicaid Services (CMS) implemented Version 3.0 of the MDS Resident Assessment Instrument (RAI). The MDS 3.0 aims to improve the validity and reliability of item response and related quality measures (QMs), increase accuracy by giving residents a voice in item responses, and enhance item instructions and quality measure specifications to better align with existing definitions and use across provider settings (Saliba & Buchanan, 2008). With the implementation of the MDS 3.0, CMS retired the QMs calculated from the earlier MDS 2.0 data. In their place, RTI has developed new nursing home QMs based on MDS 3.0 data items using the MDS 2.0 QMs as a foundation. As of July 2012, 16 MDS 3.0 quality measures are fully endorsed by the National Quality Forum (NQF). While nursing home quality data have been publicly reported for all U.S. Medicare- and Medicaid-certified nursing homes since 2002, public reporting of the new MDS 3.0 QMs began in July 2012.

This report details analytic results regarding MDS 3.0 QMs' variability, reportability, reliability, and validity. In addition, analyses examine the effectiveness of measure risk adjustment. Analyses regarding short-stay and long-stay definitions are also presented.

E.2 Data, Analytic Approach and Short and Long Population

Data

Quality measure analyses in this report were primarily based on the episode file that RTI created for calculating the QMs for the fourth quarter of 2011. The episode file primarily comprises of MDS 3.0 assessments from July 1st to December 31st (Quarters 3 and 4, 2011) but also includes some assessments from the second quarter of 2011 for some long-stay residents. The item set definition was based on the MDS 3.0 Item Listing version 1.00.2, effective October 1, 2010. Coding of the quality measures was based on the specifications detailed in the MDS 3.0 QM User's Manual, version 5.0.

Analytic Approach

This report presents quality measure analyses testing the new QMs, largely following the NQF guidance, adding a few considerations special to the MDS 3.0 measures. Five key areas, important in quality measure design, were addressed in detail:

- **Variability.** Variability refers to the QMs' ability to distinguish between high-quality and low-quality facilities, illustrating sufficient variation in quality across providers. In order to test variability, we examined the spread of the distributions of scores by calculating their interquartile ranges (the difference between the 75th and 25th percentiles) and looked for ceiling effects by calculating the proportions of facilities with perfect scores(i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes).
- **Reportability.** To be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., have at least 20 short-stay resident episodes that qualify for the denominator of the short-stay QMs and at least 30 long-stay resident episodes that qualify for the denominator of the long-stay QMs) after applying measure exclusion criteria. RTI examined the percentage of facilities for which each measure can be publicly reported, given minimum sample size requirements.
- **Reliability.** This refers to the consistency of results or the vulnerability of a measurement to random error. Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. Analyses compared facility QM scores and rank based on the QM scores across multiple quarters of data. We compared results for each facility across multiple quarters of data. We evaluated the percentage of facilities showing large (three deciles or more) changes in facility ranking (three deciles or more) and in QM scores from quarter to quarter.
- **Validity.** It is important that the QMs and individual items used to specify the QMs measure what they purport to measure, that is, the measures and items are valid. Four principal strategies were used to examine the validity of the 16 QMs.
- *Correlations*. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated).
- *Variation by State.* For a QM to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as state-level payment policies or demographics, this can be a threat to the validity of the measure. RTI examined the proportion that QM variation might be attributed to state in which the nursing home is located.
- *Missing Data.* Missing data represent a potential threat to the validity of a QM. If patterns indicate that similar types of residents and/or assessments have missing data it is likely that the measure may be not be capturing processes and outcomes for the intended population, thus inflating or suppressing facility level QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised. RTI examined the impact of missing data on QM score and the missing rate across facility. In addition, RTI examined the impact of discharge assessments on QM scores. The inclusion of

the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in one, the overall impact of including the discharge assessment on QM rates and two, the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on the QMs by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments.

- *Seasonal Variation.* If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, the measure's validity is suspect and likely impacted by factors outside of the nursing home's control. RTI examined variation in mean and median scores for each quality measure for each quarter in 2011.
- **Risk Adjustment.** Similar to the risk-adjustment methods employed with the MDS 2.0 QMs, all but two MDS 3.0 QMs use exclusion or sample restriction; indirect standardization was used for three QMs. Exclusion or sample restriction almost always include restrictions based on data availability and sometimes include sample exclusions related to risk factors or residents' clinical status. Analyses investigated variation in resident characteristics which included: (1) denominator exclusions and (2) covariates. For the QMs that risk-adjustment models are applied in the calculation of the QM score, RTI examined the impact of the selection of covariates on QM score, changes in the impact of covariates on the QM scores across quarter and the impact of risk-adjustment model specifications on the QM scores.

Short-Stay and Long-Stay Population

The publicly reported MDS 3.0 quality measures refer to distinct resident populations in nursing facilities, the short- and long-stay residents (formerly "post-acute care residents" and "chronic care residents"). The distinction between the two populations and related QMs s was intended to capture important differences in clinical and service needs between acute short-stay residents and chronic long-stay residents. The MDS 3.0 QM specification included a change in the definition of short and long-stay populations. In the MDS 2.0 QM specifications, residents were included in the post-acute care measures if they had a 14-day PPS MDS in the target quarters, and measures were calculated based only on 14-day PPS assessments. MDS 2.0 chronic care residents were identified using a quarterly or annual assessment. The analogous MDS 3.0 QM population for post acute is short-stay residents with 100 or fewer cumulative days in facility, whose assessments may be discharge assessments; 5-, 14-, 30-, 60-, 90-day PPS assessments; or admission, quarterly, annual or significant change/correction OBRA assessments. The analogous MDS 2.0 chronic care population is the MDS 3.0 long-stay population (residents whose cumulative days in the facility are greater than 100 days). In addition to clinical and service needs between the two populations, the 100-day demarcation was selected to align with the Medicare Part A benefit period—the major source of reimbursement for short stays. Out of the 16 NQF-endorsed MDS 3.0 QMs, four QM specifications are based on the short-stay population and 12 QM specifications are based on the long-stay population.

In the MDS 3.0 episode file for Quarter 4, 2011, slightly more than half of the resident episodes are classified as short stay compared to long stay (53.4 percent vs.46.6 percent). The majority of short-stay residents' target assessments are discharge return unanticipated (65 percent), followed by PPS assessment (23.8 percent). The long-stay residents' assessments tend to be quarterly assessments (64.3 percent) or comprehensive assessments (20.8 percent).

At the facility-level, the mean proportion of short-stay residents within nursing homes is 45.5 percent; the median is 42.5 percent. Some facilities tend to provide services largely to short-stay residents, while some provide services largely to long-stay residents; however, most provide services to a mix of short-stay and long-stay residents. The average length of stay is 30.1 days for short-stay residents (median = 22 days) and 433.3 days for long-stay residents (median = 485 days). Short-stay and long-stay residents have somewhat different demographic and health characteristics (e.g., a higher proportion of long-stay residents are cognitively impaired or depressed compared to short-stay residents) as well as overall expectations of care. The majority of short-stay residents are expected to be discharged to the community, whereas the majority of long-stay residents are expected to remain in the facility.

The possibility of short-stay and long-stay population misclassification was examined. About 10.6 percent of residents classified as short stay at the end of quarter 3 of 2011 accumulated sufficient days in facility during Quarter 4 of 2011to be reclassified as long stay in subsequent QM analyses. Analyses regarding the impact of these "Early Long-Stay" residents on short-stay quality measure scores indicate little change in the four short-stay QM scores if early long-stay residents are included versus excluded in the short-stay population. For all four short-stay QMs removing early long-stay residents resulted in changes of less than 1 percentage point in the facility level mean QM score. The effect on the resident-level numerator trigger rate was almost as negligible, ranging from no change [QM #0678: Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay)]) to 0.8 percentage point increase (including early long-stay residents) for QM#0676 [Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)].

E.3 Key Findings from Testing

Variability

To assess variability, we examined distributions of facility-level QM scores, focusing on the magnitude of interquartile ranges and proportions of perfect scores. Key findings include:

- Variability of QM scores is highest for the two short-stay vaccination QMs [#0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) and QM #0682 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)] and QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay).
- Variability of QM scores is particularly narrow for the six QMs with low prevalence rates.
- The highest proportions of perfect scores were for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) (50.5 percent), QM #0683 Percent of

Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) (39.4 percent), and QM #0678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay) (29.3 percent). These measures also show narrow interquartile ranges (3.2 percent, 6.4 percent, and 2.9 percent, respectively).

- Fewer than 2 percent of facilities report perfect scores on the following measures: QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) (0.3 percent), QM #0688 Percent of Residents Whose Need for Help With Activities for Daily Living Have Increased (Long Stay) (1.0 percent), and QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) (1.4 percent). These measures also demonstrate good variability with relatively wide interquartile ranges: 24.8 percent, 11.8 percent, and 17.2 percent, respectively
- QM scores were relatively high for the four vaccination measures but still exhibit good variability.

Reportability

Reportability is measured by the percentage of facilities able to provide adequate information to calculate the quality measure. This takes into account minimum sample size requirements for public reporting (i.e., had at least 20 short-stay resident episodes that qualified for the denominator of the short-stay QMs and at least 30 long-stay resident episodes that qualified for the denominator of the long-stay QMs).

- The percent of facilities able to report QMs ranges from 42.6 percent for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) to 88.4 percent for three QMs. These three QMs are QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay), QM #0683 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay), and QM #0687 Percent of Long-Stay Residents Who Were Physically Restrained (Long Stay).
- Aside from QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) and QM # 0678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay) (which had a reportability of 68.6 percent), the remaining measures each had a rate of reportability greater than 70 percent.
- Results were fairly similar comparing the proportion of facilities able to report QMs using MDS 2.0 data (~62 percent–90 percent) (Brega, Goodrich, Nuccio, & Hittle, 2008).
- The facility mean reportability for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) was 72.4 percent using MDS 2.0 data compared to 42.6 percent using MDS 3.0 data. The reportability difference between the MDS 2.0 QM and the MDS 3.0 QM is likely due to a definition change in the MDS 3.0 QM: the MDS 3.0 QM restricts target assessment to including only low-risk residents.

- We also assessed the impact of discharge assessments on reportability. Two types of analyses were performed: (1) distributions of facility level QM scores calculating the QM with and without discharge assessments in the set of target assessment and (2) changes in facility rates of reportability when discharge assessments are not included in the set of target assessments. When discharge assessments were excluded in constructing the QM, reportability was largely unchanged for the long-stay QMs. Reportability actually increased for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) by 5.3 percentage points, owing to a high rate of missing items (and thus exclusions) on pain-related items on discharge assessments. However, the loss in reportability after excluding discharge assessment are substantial for QM #678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay). Discharge assessments helped 15.6 percent of facilities report that could not without the discharge assessment. Dropping the discharge assessment may create a situation in which so many short-stay residents only have a single assessment in their episode and would be ineligible for this QM.

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decay in quality. Analyses indicate most facilities had the same approximate QM scores or slightly improved from Quarter 1, 2011, to Quarter 4, 2011. The percentage of facilities that had the same score from quarter to quarter (within one standard deviation) ranged from 71.2 percent for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) to 88.6 percent for QM #0683 Percent of Long Stay Residents Assessed and Appropriately Given the Pneumococcal Vaccine

Validity

Four principal strategies were used to examine the validity of the 16 quality measures. Key findings include:

- *Correlations*. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated). Eleven QMs were categorized into four groups (long-stay/short-stay pain; long-stay/short-stay pressure ulcer; function QMs; and four vaccination QMs). All of the measure pairs were significantly correlated except for the pairing of QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) and QM #0686, Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay). All correlations were moderate or weak between a facility's percentile ranking for one QM with its percentile ranking with another QM in the same care process group. This could be due to the differences between the short-stay and long-stay populations. In general, the correlations were weaker once the quality measure scores were risk adjusted.

- *Variation by State.* To explore whether state characteristics might be a source of facility score variation we conducted a one-way analysis of variance (ANOVA) and reported the proportion of variance in QM scores accounted for by state. There was a significant effect of state for each of the 16 quality measures. However, state location accounted for less than 5 percent of the variance in QM scores for 8 measures and greater than 10 percent for just 2: QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) at 16.2 percent and QM #0690 Percent of Residents Who Have Depressive Symptoms at 11.2 percent. We also examined the interquartile difference between the mean state-level scores for states. The proportion of variance in each QM explained by the state that facilities are located varies across QMs, ranging from 2.0 percent for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) to 16.2 percent for QM #0685 Percent of Low-risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay).
- *Missing Data.* The impact of missing data was minimal for most measures. The correlations between missing-data rates and quality measure scores tend to be weak. They also tend to indicate a direct relationship between quality measure scores and complete data keeping: facilities with higher rates of missing data tend to have poorer scores (higher scores on negative measures, lower scores on positive measures). For pain-related measures, missing data rates tended to be higher for residents who had difficulty making themselves understood, indicating a specific threat to validity posed to measures requiring residents to self-report. In addition, RTI examined the impact of discharge assessments on QM scores by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. For all measures but one, removing the discharge assessments from analyses led to a minimal reduction of the number of facilities able to report: for 14 of the 16 measures, this change represented 2 percent or fewer of the total number of facilities. Reportability actually increased after excluding discharge assessments by for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) by 5.3 percentage points, owing to a high rate of missing items (and thus exclusions) on pain-related items on discharge assessments. However, the gains in reportability are substantial for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay). Discharge assessments helped 15.6 percent of facilities report that could not without the discharge assessment.
- *Seasonal Variation.* If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, the measure's validity is suspect and likely impacted by factors outside of the nursing home's control. RTI examined variation in mean and median scores for each quality measure for each quarter in 2011 and found no evidence to suggest that seasonal variation may post a threat to validity. The widest variances in scores from quarter to quarter were seen for the two influenza vaccine measures: for each, the range between the highest and lowest mean quarter scores was about 6 percentage points. However, these results are based on just one cycle of seasons and should be considered preliminary.

Risk-Adjustment Analyses

Risk adjustment seeks to minimize the influence of factors beyond the control of nursing facilities on QM scores, maximizing the likelihood that the QM accurately assesses facility quality. Risk adjustment may also be employed to ensure that measures compare care for patients at similar risk across facilities. For the 16 NQF-endorsed QMs discussed in this report, there are two primary methods of risk adjustment: model-based adjustment and denominator exclusion or sample.

- Nine of the 16 QMs apply exclusion criteria as a method of risk adjustment. The QMs include: #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay); #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay); #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay); #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay); #0680 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay); #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay); #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay); #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay); and #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay). The most common types of exclusions include: the MDS 3.0 item was correctly skipped; the condition was not present or not assessed; the resident was not in facility during current or most recent influenza season; and if the assessment does not contain a useable response.
- A model-based risk-adjustment approach is applied in three of the 16 QMs. In this approach, MDS items, indicating certain health conditions that increase or decrease the likelihood of a health outcome, are identified as covariates. A logistic regression model is fitted and the estimated coefficients are used to predict the probability that a resident will experience that outcome given the covariate values. The average of these resident-level probabilities for a given facility represents that facility's expected score for that measure. The final QM score for a facility is found by combining its observed score (i.e., the prevalence or incidence of the outcome) with its expected score. Three QMs are risk adjusted using a model based approach. They are: QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay); QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay); and QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay). The number and covariate item differ by QM. For example, QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) incorporates four items: bed mobility, bowel incontinence, diabetes/peripheral vascular disease, and body mass index (BMI), whereas QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) applies only one item (i.e., daily decision making).
- For the QMs that risk-adjustment models are applied in the calculation of the QM score, RTI examined the impact of the selection of covariates on QM score, changes in the impact of covariates on the QM scores across quarter and the impact of risk-adjustment model specifications on the QM scores. The results suggest that the

current risk-adjustment models show satisfactory predictive power, but using hierarchical model specifications can future improve the predictive power. The selection of model specification had sizable effect on the risk adjusted QM scores and facility rank based on these QM scores.

Quality Measure–Specific Analyses

In Section 4 of this report, we present analyses for each of the 16 QMs.

QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) reports the percentage of short-stay residents who self-report daily pain with at least one episode of moderate/severe pain or very severe/horrible pain of any frequency. We present descriptive analyses and test results for this QM. Of all short-stay resident episodes, 71.5 percent meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 70.3 percent of the facilities are able to report this QM. Among facilities able to report, the mean facility QM score is 23.1 percent. The QM score varies across facilities with a standard deviation of 12.6 percent and interquartile range of 17.2 percent, suggesting acceptable variability in the measure's ability to differentiate among facilities with poor and good quality of care.

The introduction of the discharge assessment and its inclusion as a target assessment for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) is a new feature of MDS 3.0 and MDS 3.0 quality measures. Our analyses indicated that the items used to calculate QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) often have missing values on the discharge assessment. Assessments with missing item values, and thus resident episodes, are dropped from the calculation of the measure, affecting QM reportability. Analyses illustrated that including discharge assessments in the set of target assessments decreases the resident episodes in the denominator by about 8 percent and the percent of facilities able to report this QM by approximately 5 percent.

As reliability checks of this QM, we tracked changes in QM score and in percentile ranking by quarter. The majority of facility score changes varied only one standard deviation, with a very low proportion of facility QM score changes greater than three standard deviations. A similar pattern was found for facility rank changes. About two-thirds of facilities remained within the same decile from quarter to quarter. Almost all facilities had rank changes within three deciles. These findings demonstrate acceptable reliability for this QM.

Several validity tests were conducted for this QM. First, the correlations between QMs within a "care process group" were analyzed. The hypothesis was that changes in one QM within a measure group (e.g., short-stay and long-stay pain QMs) should be correlated with changes in other measures because they reflect similar care processes. Findings illustrated statistically significant and moderate correlations between the short-stay and long-stay pain QMs ($r = 0.560$; $p < .001$). Second, we examined the frequency and distribution of missing data. We also found that missing rate only varies slightly by resident characteristics (e.g., age, gender, and cognitive status). The missing rate also varies across facilities and was higher among facilities

with higher values in the QM score (indicating poorer quality). Finally, we analyzed the potential geographic (state) and seasonal (quarterly) variations in this QM. The state in which the facility is located explains about 7 percent of the variation in this QM. The QM score remains largely stable from Quarter 1 to Quarter 4 in 2011.

Reliability and validity tests indicate acceptable rigor. QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) received full NQF endorsement on August 1, 2012.

QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) reports the percentage of short-stay residents who had one or more new or worsening Stage 2 to 4 pressure ulcers during the target episode. This QM represents an improvement on prior MDS 2.0 measure because it focuses on pressure ulcers that develop within the nursing facility and excludes Stage 1 ulcers. This is an incidence measure and requires that the resident have at least one nonadmission target assessment. This QM is risk adjusted using indirect standardization, adjusting for four covariates based on the resident's initial assessment in the episode: the need for assistance in bed mobility self-performance, occasional bowel incontinence, the presence of a diabetes or peripheral vascular disease, and low BMI.

We found that 65.0 percent of the short-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, a little more than two thirds of the facilities are able to report the risk adjusted QM. The need for a qualifying initial assessment to obtain the covariate values had a large impact on the sample size and reportability. In the cases in which no initial assessment was available, an episode was not included in the QM definition. About 28 percent of the short-stay residents did not have a qualifying initial assessment. Additionally, 6.6 percent have an initial assessment but have missing values for the covariates.

In general this was a low incidence QM (mean score: risk adjusted 1.9 percent, unadjusted 1.9 percent). Approximately one-third of facilities (35.4 percent for the risk-adjusted QM, 36.2 percent for the unadjusted) had no residents with new or worsened pressure ulcers. The QM score varied across facilities with a standard deviation of 2.4 percent (2.3 percent for the unadjusted measure) and interquartile range of 2.9 percent (2.8 percent for the unadjusted measure).

To examine the reliability of this QM, we analyzed the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. About half of facilities improved by less than one standard deviation and 26.7 percent of facilities decreased in their score by less than one standard deviation. A very low proportion of facilities' scores changed over three standard deviations. About 60 percent of facilities remained within the same decile rank from quarter to quarter. In the first three quarters of 2011, about one-fifth of facilities hasd rank changes of more than three deciles. From Quarter 3 to Quarter 4, only 6.5 percent of facilities shift more than three deciles in ranking. The relatively large rank changes in some quarters may be caused by the low incidence rate for this QM.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." It was hypothesized that changes in one QM within a measure group (e.g., short-stay and long-stay pressure ulcer QMs) should be correlated with changes in other measures because they reflect similar care processes. The findings from this analysis showed significant but weak correlations among the short- and long-stay pressure ulcer measures ($r = 0.148$). Second, we examined the frequency and distribution of missing data and found that 6.6 percent of the short-stay population was excluded because they were missing data on the items used to calculate BMI. The missing rate varies across facilities but there is little evidence to indicate a substantial relationship with the QM scores. This indicates that missing data does not pose a threat to QM validity. Finally, we analyzed the potential geographic (state) and seasonal (quarterly) variations in this QM. There was little evidence for either. The state in which the facility is located explained only about 2 percent of the variation in this QM. Regarding seasonal variation, the QM score declined only slightly from Quarter 1 to Quarter 4 in 2011.

This QM uses model-based risk adjustment of four covariates. We examined the impact of partial risk adjustment (adjusting for subsets of the four covariates) on the risk-adjusted QM scores and facility score changes. We also compared coefficients for the covariates across quarters. The results support the validity of the risk-adjustment model. In addition, we explored different model specifications (single-level logistic regression models vs. hierarchical models) for risk adjustment. The results show that changing model specifications has a large impact on facility rank based on the risk-adjusted QM score.

Reliability and validity tests indicate acceptable rigor. QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) received full NQF endorsement on August 1, 2012.

QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay)

QM#0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay) reports the percentage of short-stay residents who either received the influenza vaccine during the current or most recent influenza season (either in the facility or outside the facility) or were offered and declined the vaccine or were ineligible due to contraindication(s). This is an important measure of quality of care in the nursing facility, as morbidity and mortality related to influenza are often reported in conjunction with data regarding pneumonia, and together frequently lead to death in the elderly population.

About 71.7 percent of the short-stay resident episodes meet the denominator inclusion criteria and were included in the denominator. Based on these resident episodes, 75.4 percent of facilities have 20 or more short-stay episodes included in the denominator and are able to report this QM. Among facilities able to report, the mean facility QM score is 79.7 percent. The QM score varies across facilities with a standard deviation of 19.2 percent and an interquartile range of 23.3 percent. This indicates acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To examine the reliability of this QM, changes in QM score and in rank based on the QM score for each facility were examined by quarter. About half of facilities declined in their score by less than one standard deviation and 26.7 percent of facilities decreased in their score by less than one standard deviation. In a very low proportion of facilities score changes were greater than three standard deviations. A similar pattern was found for facility percentile ranking changes. About two-thirds of facilities remained within the same decile from quarter to quarter. Almost all facilities had rank changes within three deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group (e.g., vaccination QMs) should be correlated with changes in other measures because they reflect similar care processes. Findings from this analysis showed moderate to high correlations among the vaccination QMs. Given that values of zero are imputed for missing data on items used to calculate this quality measure, no resident episode was excluded due to missing data. We examined the imputation rate across facilities and found very weak relationship between imputation rate and facility QM score. The result suggests that missing data should not be a threat to the validity of this QM. We also analyzed the potential geographic (state) and seasonal (quarterly) variations in this QM. The state in which the facility is located explains about 4 percent of the variation in this QM. The QM score slightly fluctuates across quarters in 2011 and peaks in Quarter 2, 2011.

QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)

QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) reports the percentage of short-stay residents whose pneumococcal vaccination is up to date or who were offered and declined the vaccine or were ineligible due to contraindication(s). As this QM does not have denominator exclusions, all short-stay residents are included in the denominator. Almost all facilities (99.7 percent) have 20 or more short-stay resident episodes included in the denominator and are able to report this QM. Among facilities able to report, the mean facility QM score is 81.0 percent. The QM score varies across facilities with a standard deviation of 21.0 percent and interquartile range of 23.0 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To examine the reliability of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. About half of facilities declined in their score by less than one standard deviation and 31.2 percent of facilities increased their score by less than one standard deviation. A very low proportion of changes are greater than three standard deviations. A similar pattern was found for facility rank changes. About 80 percent of facilities remained within the same decile from quarter to quarter. Only about 2 percent of facilities shift more than three deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within

a measure group (e.g., vaccination QMs) should be correlated with changes in other measures because they reflect similar care processes. Findings from this analysis showed that this QM is highly correlated with the long-stay pneumococcal vaccination QM ($r = 0.690$) and the short-stay influenza vaccination QM ($r = 0.732$). This QM is statistically significantly but moderately correlated with the long-stay influenza vaccination QM ($r = 0.342$; $p <.001$). We also analyzed the potential geographic (state) and seasonal (quarterly) variations in this QM. The state of the facility explains about 5 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)

QM#0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) reports the percentage of long-stay residents who have one or more falls that resulted in major injury during the reporting period. We found that all long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 88.4 percent of the facilities have 30 or more long-stay resident episodes included in the denominator and are able to report this QM. Among facilities able to report, the mean facility QM score is 3.4 percent. The QM score varies across facility with a standard deviation of 2.7 percent and interquartile range of 3.3 percent. The relatively small variability of this QM is largely due to the low incidence rate.

To determine the reliability of this QM, we examined changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A similar pattern was found for facility rank changes. About 60 percent of facilities remained within the same decile from quarter to quarter. About 90 percent of facilities have rank changes within three deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures because they reflect similar care processes. We selected QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the two QMs (correlation $r = 0.111$). Second, we examined the frequency and distribution of missing data. Missing rate is very low (nearly zero) for the items used to construct this QM and therefore should not pose a threat to validity of the QM. Last but not the least, we also analyzed the potential geographic (state) variations in this QM. The state in which a facility is located explains about 7.8 percent of the variation in this QM. RTI will examine this issue further when more data becomes available.

Reliability and validity tests indicate acceptable rigor. QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) received full NQF endorsement on August 1, 2012.

QM #0677 Percentage of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) reports the percentage of long-stay residents who self-report daily pain and at least one episode of moderate/severe pain or very severe/horrible pain of any frequency. This QM is risk adjusted using indirect standardization, adjusting for one covariate based on the resident's prior assessment in the episode. The covariate has a value of 1 if the resident shows independence or modified independence in daily decision making (MDS 3.0 item C1000 = [0,1]) or has a BIMS score from 13 to 15.

We found that 71.8 percent of the long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, more than 70 percent of the facilities have 30 or more long-stay resident episodes included in the denominator and are able to report this QM. Among facilities able to report, the mean facility risk-adjusted QM score is 11.5 percent. The QM score varies across facilities with a standard deviation of 8.5 percent and interquartile range of 11.5 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To determine the reliability of this QM, we examined the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. About 60 percent of facilities remained within the same decile from quarter to quarter. About 7 percent of facilities had rank changes of more than three deciles from quarter to quarter in 2011.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group (e.g., short-stay and long-stay pain QMs) should be correlated with changes in other measures because they reflect similar care processes. The findings from this analysis showed significant and moderate correlations between the short-stay and long-stay pain QMs (correlation $r = 0.560$). Second, we examined the frequency and distribution of missing data. We found that missing rate varies by some resident characteristics (e.g., residents 85 years old or older) or residents with cognitive impairment were more likely to have missing data. The correlation analysis shows that the correlation between the missing rate and the QM score is very weak, indicating that missing data should not pose a threat to validity of the QM. Finally, we also analyzed the potential geographic (state) and seasonal (quarter) variations in this QM. The state in which the facility is located explains about 9.5 percent of the variation in this QM. The QM score remains largely stable from Quarter 1 to Quarter 4 in 2011.

This QM uses model-based risk adjustment adjusting for one covariate, which is independence or modified independence in daily decision making on the prior assessment. We examined the impact of risk adjustment on the QM scores and facility score changes. We also compared coefficients for the covariates across quarter. The results support the validity of the risk-adjustment model. In addition, we explored different model specifications (single-level logistic regression models vs. hierarchical models) for the risk-adjustment model. The results

show that changing model specifications has a big impact on facility rank based on the risk-adjusted QM score.

Reliability and validity tests indicate acceptable rigor. QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) received full NQF endorsement on August 1, 2012.

QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) reports the percentage of long-stay residents at a high risk for pressure ulcers who have Stage 2 to 4 pressure ulcers. This QM is risk adjusted using denominator exclusions. Residents not at a high risk for pressure ulcers—defined as impaired bed mobility or transfer, comatose or malnutrition or at risk of malnutrition—are excluded from the denominator.

We found that 67.8 percent of the long-stay resident episode meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 72.4 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 6.9 percent. The QM score varies across facility with a standard deviation of 4.6 percent and interquartile range of 6.2 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To examine the reliability of this QM, we analyzed the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. About half of facilities remained within the same decile from quarter to quarter. About 17 percent of facilities have rank changes of more than three deciles from quarter to quarter. Compared with other QMs, the relatively large rank changes may be due to the restrictive denominator inclusion criteria (i.e., high risk) and thus smaller denominator size.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group (e.g., short-stay and long-stay pressure ulcers QMs) should be correlated with changes in related measures because they reflect similar care processes. Findings from this analysis showed significant but weak correlations among the short-stay and long-stay pressure ulcer measures (correlation $r = 0.148$). The weak correlation may be due to the different specifications for the short-stay and long-stay QMs (i.e., the short-stay QM is an incidence measure, whereas the long-stay QM is a prevalence measure). It may also suggest that the care processes for preventing pressure ulcers and/or the patient characteristics in the short-stay versus long-stay populations are very different, and that facilities may have more specialized expertise in dealing with only one of the patient groups. Second, we examined the frequency and distribution of missing data. The rate of missing data is very low (nearly zero) for items used to calculate this QM and therefore should not pose a threat to validity of the QM. Finally, we analyzed the potential geographic (state) and seasonal (quarter) variations in this QM. The state

in which the facility is located explains about 6.5 percent of the variation in this QM. The QM score remains largely stable from Quarter 1 to Quarter 4 in 2011.

Reliability and validity tests indicate acceptable rigor. QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) received full NQF endorsement on August 1, 2012.

QM #0681 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Long-Stay)

QM #0681 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Long Stay) reports the percentage of long-stay residents who either received the influenza vaccine during the current or most recent influenza season (either in the facility or outside the facilities) or offered and declined the vaccine or were ineligible due to contraindication(s). This is an important measure of quality of care in the nursing facility, as morbidity and mortality related to influenza are often reported in conjunction with data regarding pneumonia, and together frequently lead to death in the elderly population.

We present descriptive analyses and test results for this QM. We found that 98.5 percent of the long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 88 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 89.1 percent. The QM score varies across facility with a standard deviation of 13.7 percent and interquartile range of 14.8 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To check the reliability of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. For about 44 percent of the facilities, the QM score increased by less than one standard deviation. For 38.5 percent of facilities, the QM score decreased by less than one standard deviation. From Quarter 3 to Quarter 4, 2011 about half of facilities remained within the same decile, but 17.6 percent if facilities have rank changes of more than three deciles.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group (e.g., vaccination QMs) should be correlated with changes in other measures because they reflect similar care processes. Findings from this analysis showed moderate to high correlations among the vaccination QMs. We also analyzed the potential geographic (state) and seasonal (quarter) variations in this QM. The state in which the facility is located explains about 3.3 percent of the variation in this QM. The QM score slightly declined from Quarter 1 to Quarter 4 in 2011.

QM #0683 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)

QM #0683 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay) reports the percentage of long-stay residents whose pneumococcal vaccination is up to date or who were offered and declined the vaccine or were

ineligible due to contraindication(s). As this QM does not have denominator exclusions, all short-stay residents are included in the denominator. About 88 percent have 30 or more long-stay resident episodes included in the denominator and are able to report this QM. Among facilities able to report, the mean facility QM score is 93.8 percent. The QM score varies across facility with a standard deviation of 12.4 percent and interquartile range of 6.4 percent. About 40 percent of facilities have "perfect scores" (i.e., 100 percent). These findings indicate that most facilities perform well with regard to this aspect of care. The relatively large standard deviation suggests that this QM can be particularly useful to identify facilities with poor quality.

To determine the reliability of this QM, we examined the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The large majority of facility score changes, either improvements or declines, were within one standard deviation. A very low proportion of changes were greater than three standard deviations. A similar pattern was found for facility rank changes. More than 80 percent of facilities remained within the same decile from quarter to quarter. Only about 2 percent of facilities shift more than three deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group (e.g., vaccination QMs) should be correlated with changes in other measures because they reflect similar care processes. Findings from this analysis showed that this QM is highly correlated with the short-stay pneumococcal vaccination QM (correlation $r = 0.690$) and the short-stay influenza vaccination QM (correlation $r = 0.533$). This QM is significantly but moderately correlated with the long-stay influenza vaccination QM (correlation $r = 0.392$). We also analyzed the potential geographic (state) and seasonal (quarter) variations in this QM. The state in which the facility is located explains about 4.5 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

QM #0684 Percent of Residents with a Urinary Tract Infection (Long-Stay)

QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) reports the percentage of long-stay residents whose target assessment indicating urinary tract infection within the last 30 days. We present descriptive analyses and test results for this QM. We found that 98.1 percent of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 87.7 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 7.7 percent. The QM score varies across facility with a standard deviation of 5.7 percent and interquartile range of 7.3 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

As reliability checks of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. About half of facilities remained within the same decile from quarter to quarter. About 90 percent of facilities are with rank changes within three deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group". The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures because they reflect similar care processes. We selected QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) and QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the three QMs. Second, we examined the frequency and distribution of missing data. Missing rate is very low for the items used to construct this QM and therefore should not pose a threat to validity of the QM. Lastly, we analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 3.8 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)

QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) reports the percentage of low-risk long-stay residents whose target assessment indicating frequently or always incontinence of the bladder. This QM is risk adjusted applying denominator exclusions. Residents at high risk for incontinence are excluded from the denominator of this quality measure. High risk is defined as any of the following conditions: severe cognitive impairment; totally dependent in bed mobility, transfer, or locomotion.

We present descriptive analyses and test results for this QM. Of long-stay resident episodes, 39.7 percent meet the denominator inclusion criteria and are thus included in the denominator. Based on these resident episodes, 42.7 percent of the facilities are able to report this QM (i.e., facilities with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 42.6 percent. The QM score varies across facility with a standard deviation of 17.3 percent and interquartile range of 35.2 percent, indicating acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To measure, we assessed changes in QM score and in rank based on the QM score by facility by quarter. The majority of facility score changes (either improvements or declines) were within one standard deviation. A very low proportion of changes being greater than three standard deviations. A similar pattern was found for facility rank changes. Most facilities (about 90 percent) illustrate rank changes within three deciles.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures since the measures reflect similar care processes. We selected QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) and QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the three QMs. Second, we examined the frequency and distribution of missing data. On average, 1 percent of resident episodes cannot be used for calculating this QM due to missing data. Missing data, therefore, should not

pose a threat to validity of the QM. Last but not the least, we also analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 16.2 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) reports the percentage of long-stay residents whose target assessment indicating the use of indwelling catheters. This QM is risk adjusted based on logistic regression models, and two covariates are used: whether the resident has frequent bowel incontinence and whether the resident has pressure ulcers at stages 2 through 4 on the prior assessment.

Using MDS 3.0 data, 81.7 percent of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 86.9 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility risk adjusted QM score is 4.1 percent. The risk-adjusted QM score varies across facility with a standard deviation of 3.3 percent and an interquartile range of 4.1 percent.

To check the reliability of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A similar pattern was found for facility rank changes. Most of facilities are with rank changes within three deciles.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures because they reflect similar care processes. We selected QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) and QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the three QMs. Second, we examined the frequency and distribution of missing data. Missing rate is very low for the items used to construct this QM and therefore should not pose a threat to validity of the QM. Last but not the least, we also analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 4.0 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

This QM is risk adjusted based on logistic regression models with two covariates. We examined the impact of partial risk adjustment (risk adjusting for subsets of the four covariates) on the risk-adjusted QM scores and facility score changes. We also compared coefficients for the covariates across quarter. The results support the validity of the risk-adjustment model. In addition, we explored different model specifications (single-level logistic regression models vs. hierarchical models) for the risk-adjustment model. The results show that changing model specifications has a big impact on facility rank based on the risk-adjusted QM score, as more

than half of the facilities with extremely poor performance flagged using the hierarchical model were not flagged using the single-level model.

QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) reports the percentage of long-stay residents whose target assessment indicating daily physical restraints. We present descriptive analyses and test results for this QM. Almost all (99.9 percent) of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 88.4 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 2.4 percent. The QM score varies across facility with a standard deviation of 4.2 percent and interquartile range of 3.2 percent. The overall low-mean QM score and the relatively large variation suggest that this QM can be useful to identify facilities with poor quality of care.

To check the reliability of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A similar pattern was found for facility rank changes. About 90 percent facilities are with rank changes within three deciles.

We conducted several validity tests for this QM. First, we examined the frequency and distribution of missing data. Missing rate is very low for the items used to construct this QM and therefore should not pose a threat to validity of the QM. We also analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 6.9 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

QM #0688 Percentage of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) reports the percentage of residents whose need for help with late-loss activities of daily living (ADLs) has increased. We present descriptive analyses and test results for this QM. We found that 81.7 percent of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 81.2 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 16.7 percent. The QM score varies across facility with a standard deviation of 9.3 percent and interquartile range of 11.8 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To check the reliability of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A slightly different pattern

was found for facility rank changes. From quarter to quarter in 2011, about 20 percent of facilities are with rank changes more than three deciles.

We conducted several validity tests for this QM. QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures because they reflect similar care processes. We selected QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the two QMs (correlation $r = 0.111$). Second, we examined the frequency and distribution of missing data. Missing rate is about 1 percent for the items used to construct this QM and therefore should not pose a threat to validity of the QM. Last but not the least, we also analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 6 percent of the variation in this QM. The QM score only slightly decreased from quarter to quarter in 2011.

QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)

QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) reports the percentage of residents who had a weight loss of 5 percent or more in the last month or 10 percent or more in the last two quarters who were not on a physician prescribed weight-loss regimen. We present descriptive analyses and test results for this QM. We found that almost all (97.4 percent) of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 87.7 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 7.1 percent. The QM score varies across facility with a standard deviation of 4.6 percent and interquartile range of 5.8 percent.

To check the reliability of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A slightly different pattern was found for facility rank changes. From quarter to quarter in 2011, about 20 percent of facilities are with rank changes more than three deciles.

We conducted several validity tests for this QM. First, we examined the frequency and distribution of missing data. Missing rate is less than 1 percent for the items used to construct this QM and therefore should not pose a threat to validity of the QM. We also analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 3.3 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) reports the percentage of long-stay residents who have had symptoms of depression during the 2-week period preceding the MDS 3.0 target assessment date. We present descriptive analyses and test

results for this QM. We found that 97.5 percent of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 87.8 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 7.3 percent. The QM score varies across facility with a standard deviation of 10.8 percent and interquartile range of 8.8 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

As reliability checks of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A similar pattern was found for facility rank changes. From quarter to quarter in 2011, more than 90 percent of facilities are with rank changes within three deciles.

We conducted several validity tests for this QM. First, we examined the frequency and distribution of missing data. Missing rate is 2.5 percent for the items used to construct this QM. The missing rate varies across facility, with some facilities having a missing rate greater than 7.5 percent. However, the missing rate is not significantly associated the QM score, indicating that missing data should not pose a threat to validity of the QM. We also analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 11.2 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

Reliability and validity tests indicate acceptable rigor. QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) received full NQF endorsement on August 1, 2012.

E.4 Summary

The MDS 3.0 was implemented October 1, 2010, replacing the MDS 2.0. The MDS 3.0 aimed to improve the reliability of item response and related quality measures; increase accuracy by giving the residents voice in item responses; and enhance item and quality measure instructions to better align with existing definitions and use across provider settings.

This report summarizes the findings regarding MDS 3.0 nursing home QMs. The current publicly reported quality measures refer to distinct resident populations in nursing facilities, the short-stay and long-stay populations (formerly “post-acute care residents” and “chronic care residents”). The introduction of MDS 3.0 was accompanied by a change in the short-stay definition (100 days or fewer cumulative days in the facility) and long stay (over 100 cumulative days in the facility). One consequence of the new definition of short stay is that about 10.6 percent of short-stay residents in a target quarter may become long-stay in the next quarter as they stay in the nursing homes and eventually accumulate more than 100 days of stay (i.e., early long-stay residents in the target quarter). The impact of including these early long-stay residents in constructing short-stay QMs was examined.

Based on the MDS 3.0 episode file for Quarter 4, 2011, RTI analyzed the variability, reportability, reliability, validity, and risk adjustment of the QMs. The results suggest that the 16 quality measures perform well across analytic domains. Generally, these 16 measures have

scores which vary widely enough to discriminate between facilities with different levels of quality of care and are reliable and valid.

This page intentionally left blank.

SECTION 1
INTRODUCTION

1.1 Statement of Purpose

An estimated 1.4 million U.S. residents currently live in 15,700 nursing homes (NHs). The quality of care in these facilities has long been a source for national concern. One national effort to improve quality, included in the Omnibus Budget Reconciliation Act (OBRA) of 1987, was a national mandate to implement a standardized assessment for persons residing in NHs. The resulting assessment, the minimum data set (MDS), is a standardized instrument used to assess the health and functional status and is required for all residents and patients receiving care in Medicare and Medicaid certified NHs. The Minimum Data Set (MDS) component of the Resident Assessment Instrument was designed to facilitate achievement and maintenance of the resident's highest practical level of well-being by identifying and documenting individual resident's needs and strengths, and incorporating this information into the resident's care plan (Centers for Medicare & Medicaid Services, 2010b).

On October 1, 2010, the Centers for Medicare and Medicaid Services (CMS) implemented Version 3.0 of the Minimum Data Set (MDS) Resident Assessment Instrument (RAI). The MDS 3.0 aims to improve the validity and reliability of item response and related quality measures (QMs); increase accuracy by giving residents a voice in item responses; and enhance item instructions and quality measure specifications to better align with existing definitions and use across provider settings (Saliba & Buchanan, 2008). With the implementation of the MDS 3.0, CMS retired the QMs calculated from the earlier MDS 2.0 data. In their place, RTI has developed new nursing home QMs based on MDS 3.0 data items using the MDS 2.0 QMs as a foundation. Input regarding new QMs was solicited from technical and clinical experts, various divisions within CMS, and consumer and provider stakeholders. CMS and RTI submitted 17 MDS 3.0 QMs to the National Quality Forum (NQF) review process, and these measures were formally endorsed in February 2011. Ten measures received full endorsement, and seven received time-limited endorsement (TLE). CMS later withdrew one new TLE QM (scheduled pain regimen – short stay) from NQF endorsement, while the remaining 6 TLE QMs received full endorsement in July 2012. Public reporting on Nursing Home Compare of the new 16 MDS 3.0 QMs began in July 2012.

This report details analytic results regarding MDS 3.0 QMs reportability, reliability, and validity. In addition analyses examined the effectiveness of measure risk adjustment and the ability to discriminate among nursing home providers with varying performance.

1.2 History of MDS 3.0 QM Development

In 1986, the Institute of Medicine report, *Improving the Quality of Care in Nursing Homes,* recommended that systematic and standardized assessments of residents' cognitive, functional, and emotional needs be used to assess nursing home quality (Institute of Medicine, 1986). Shortly thereafter, the Omnibus Budget Reconciliation Act of 1987 mandated the development of a resident assessment instrument capable of measuring nursing home residents' health and quality of life. A result of this legislation was the implementation of a standardized assessment instrument, the MDS, which was designed to assess nursing home residents' functional status, mood, and medical conditions. These MDS data are used nationwide in

Medicare- and Medicaid-certified nursing homes to assess residents clinically, report nursing home quality publicly, and reimburse nursing homes for Medicare skilled nursing care. The MDS assessment was implemented nationally in 1990, was updated in 1995 and 1998, and was enhanced in 2010 (MDS version 3.0).

All Medicare- and Medicaid-certified nursing homes are required to collect and report MDS data to CMS on admission, quarterly, and annually, as well as upon a significant change in resident status, and to submit significant corrections to prior comprehensive or quarterly assessments. In addition, providers must complete assessments for payment under the Medicare Part A benefit for beneficiaries who receive post-acute care in a skilled nursing facility (SNF). These assessments are completed at 5, 14, 30, 60, and 90 days of the Medicare Part A stay and upon readmission or return to the facility. The MDS component of the Resident Assessment Instrument (RAI) was designed to facilitate the achievement and maintenance of residents' highest practical level of well-being by identifying and documenting individual residents' needs and strengths, and incorporating this information into residents' care plans—clinical elements used to monitor residents' quality of care. With the implementation of the MDS 3.0, clinical information is also being collected at discharge from the facility.

Nursing home quality data have been publicly reported for all U.S. Medicare- and Medicaid-certified nursing homes since 2002. These are used to assist the facility in monitoring and improving the quality of care for all residents. Recent changes in the MDS 3.0 assessment provided an opportunity for QM refinement, capitalizing on revised definitions and new data items (e.g., resident interviews).

In 2006, CMS launched the Data Assessment and Verification 2 project (DAVE-2) , project, the goal of which was to "assess the integrity of MDS information, and to measure and improve the accuracy of MDS assessments submitted by nursing facilities." The DAVE-2 project involved site visits by nurse reviewers to: conduct detailed reviews of MDS assessments and independent resident assessments; and provide educational support to nursing facility staff. The DAVE-2 provided important information on the accuracy, reliability, and validity of the MDS assessment, as well as identified areas for improvement.

This work on the MDS data was advanced under CMS's Development, Maintenance, and Implementation of Nursing Home Quality of Care Measures (DMINHo) project. The DMINHo was an empirical review of publicly reported nursing home quality measures, examining them for reportability, variability, stability, validity, reliability, and risk-adjustment feasibility. DMINHo conducted analyses using both national facility-level data (national QM data and facility characteristics from the Online Survey, Certification, and Reporting [OSCAR]) and resident-level data (random sample of MDS assessments, using findings from the DAVE-2 project). The DMINHo project determined that none of the quality measures had acceptable performance in all six areas, and two measures had unacceptable performance in all measures (Delirium and Low-Risk Pressure Ulcers). These findings identified the areas that needed improvement and refinement, providing the basis for the NHQM work on updating the measures.

In summer 2009, RTI met with CMS staff to consider topic areas for MDS 3.0 QM development. RTI submitted a report to CMS in September 2009 that summarized recommendations from CMS and information about MDS 3.0 nursing home QMs developed for

the NQF endorsement process in 2010 and provided information about additional high-priority potential measures for CMS review. A Technical Expert Panel (TEP) was convened in October 2009 to solicit input regarding potential new measures. The TEP voted to move forward immediately with two new MDS 3.0 QMs focusing on falls and on pain management for submission to NQF in March 2010 along with 15 other QMs based on MDS 2.0 implementation experience and related research (Saliba & Buchanan, 2008).

Table 1-1 provides the list of QMs submitted to and endorsed by NQF. The table includes key defining characteristics of the measures and also indicates the 10 measures that received full endorsement at the time of initial review, the seven that received time-limited endorsement, including the TLE QM that was subsequently withdrawn by CMS from NQF consideration.[1] For the TLE measures, NQF required that the CMS and RTI submit reliability and validity testing results in February 2012, 12 months after the initial limited endorsement was granted. NQF required additional testing results for the TLE QMs because they reflected one or more of the following characteristics: (1) the measure is newly developed; (2) the underlying items used to calculate the QM changed significantly from MDS 2.0 to MDS 3.0; or (3) the QM, as a construct, substantially changed during the transition from MDS 2.0 to MDS 3.0.

[1] Additionally, two of the fully endorsed quality measures (short-stay pneumococcal and influenza vaccination quality measures) were expanded to include two post-acute care providers—long-term care hospitals (LTCHs) and inpatient rehabilitation facilities (IRFs). NQF "opened" the two quality measures in Fall 2011 for expansion with ad hoc review and the QMs received NQF endorsement for expansion in May 2012. A third measure, the short-stay pressure ulcer measure, was reviewed for expansion to LTCHs concurrently with the TLE measure review in July 2012 and received full endorsement.

Table 1-1
CMS NQF endorsed quality measures: names and key characteristics

NQF #	Quality measure title	Time limited (TLE)	Measure type	Covariate adjustment	Seasonal variation adjustment	Direction indicating improvement
	SHORT-STAY QUALITY MEASURES					
#0675	The Percent of Residents on a Scheduled Pain Medication Regimen on Admission Who Self-Report a Decrease in Pain Intensity or Frequency (Short Stay)	NEW TLE Withdrawn	Incidence	—	—	Increase OUTCOME QM
#0676	Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)	TLE 5 Star	Prevalence	—	—	Decrease OUTCOME QM
#0678 (expanded to LTCH & IRF)	Percent of Residents With Pressure Ulcers That are New or Worsened (Short Stay)	TLE 5 Star	Incidence	X (from initial assessment)	—	Decrease OUTCOME QM
#0680 (expanded to LTCH & IRF)	Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)	—	Prevalence	—	—	Increase PROCESS QM
#0682 (expanded to LTCH & IRF)	Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay)	—	Prevalence	—	—	Increase PROCESS QM
	LONG-STAY QUALITY MEASURES					
#0674	Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)	NEW TLE 5 Star	Incidence	—	X	Decrease OUTCOME QM
#0677	Percent of Residents who Self-Report Moderate to Severe Pain (Long Stay)	TLE 5 Star	Prevalence	X (from prior assessment)	—	Decrease OUTCOME QM
#0679	Percent of High-Risk Residents With Pressure Ulcers (Long Stay)	TLE 5 Star	Prevalence	—	—	Decrease OUTCOME QM

(continued)

Table 1-1 (continued)
CMS NQF endorsed quality measures: Names and key characteristics

NQF #	Quality measure title	Time limited (TLE)	Measure type	Covariate adjustment	Seasonal variation adjustment	Direction indicating improvement
#0681	Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)	—	Prevalence	—	—	Increase PROCESS QM
#0683	Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)	—	Prevalence	—	—	Increase PROCESS QM
#0684	Percent of Residents with a Urinary Tract Infection (Long Stay)	5 Star	Prevalence	—	X	Decrease OUTCOME QM
#0685	Percent of Low Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)	—	Prevalence	—	—	Decrease OUTCOME QM
#0686	Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)	5 Star	Prevalence	X (from prior assessment)	—	Decrease PROCESS QM
#0687	Percent of Residents Who Were Physically Restrained (Long Stay)	5 Star	Prevalence	—	—	Decrease PROCESS QM
#0688	Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)	5 Star	Incidence	—	—	Decrease OUTCOME QM
#0689	Percent of Residents Who Lose Too Much Weight (Long Stay)	—	Prevalence	—	X	Decrease OUTCOME QM
#0690	Percent of Residents Who Have Depressive Symptoms (Long Stay)	TLE	Prevalence	—	—	Decrease OUTCOME QM

NOTE: #0675 The Percentage of Residents on a Scheduled Pain Medication Regimen on Admission Who Self-report a Decrease in Pain Intensity or Frequency (Short Stay) was withdrawn from NQF review process due to low reportability.

1.3 Analytic Approach and Report Organization

1.3.1 Overview

In June 2010, NQF provided direction to measure developers about their review criteria, in their publication: *Review of Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties, Draft Report for Review and Comment* (National Quality Forum, 2010). NQF's evaluation ratings for reliability and validity are high, moderate and low. The analyses summarized in this report provide evidence to meet the "high" evaluation rating as described by NQF (see ***Table 1-2*** for high and moderate rating criteria). For several aspects of measure testing, RTI also followed the approach used by the University of Colorado to evaluate the previously endorsed MDS 2.0 QMs (Brega, Goodrich, Nuccio, et al., 2008). In the following sections we present analyses related to MDS 3.0 QMs' reportability, variability, stability, validity and risk-adjustment models. The reliability of the items used to construct the current set of QMs has already been tested by evaluating the reliability of the MDS 3.0 item set conducted by RAND Corporation (Saliba & Buchanan, 2008). Section 2 examines the new definitions of short-stay and long-stay residents that replaced the MDS 2.0 chronic and post-acute populations, an important conceptual change in the QM definition. Sections 3 and 4 discuss further analyses addressing reliability and comparisons with MDS 2.0 QMs.

Table 1-2
NQF evaluation criteria for reliability and validity*

Rating	Reliability	Validity
High	All measure specifications (e.g., numerator, denominator, exclusions, risk factors, scoring) are unambiguous and likely to consistently identify who is included and excluded from the target population and the event, condition, or outcome being measured; how to compute the score, etc.; **AND** Empirical evidence of reliability of both data elements and measure score: Data element reliability statistics for critical data elements and measure score are within acceptable norms (tested, or reported in the literature for the same data source); **OR** commonly used data elements with little question of reliability (e.g., gender, age, date of admission); **OR** *may forego data element reliability testing if data element validity demonstrated*; **AND** Measure score reliability (precision) statistic within acceptable norms.	The measure specifications (numerator, denominator, exclusions, risk factors) reflect the quality of care problem (1a,1b) and evidence cited in support of the measure focus (1c) under *Importance to Measure and Report*; **AND** Empirical evidence of validity of both data elements and measure score: Data element validity statistical testing results are within acceptable norms; **AND** Measure score validity testing demonstrates a statistically significant result for the hypothesized performance of the measure score; **AND** Identified threats to validity (lack of risk adjustment/stratification, multiple data sources/methods, systematic missing or "incorrect" data, statistical methods) are empirically assessed and adequately addressed in measure specifications.

(continued)

Table 1-2 (continued)
Evaluation ratings for reliability and validity*

Rating	Reliability	Validity
Moderate	All measure specifications are unambiguous as noted previously; **AND** Empirical evidence of reliability for either data elements **OR** measure score as noted previously.	The measure specifications reflect the evidence cited under *Importance to Measure and Report* as noted previously; **AND** Empirical evidence of validity for either data elements **OR** measure score as noted above; **OR** Systematic assessment of face validity of measure score as a quality indicator explicitly addressed and found substantial agreement that ***the scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good and poor quality*** **AND** Identified threats to validity noted above are empirically assessed and adequately addressed in measure specifications.

SOURCE: National Quality Forum, 2010.

1.3.2 Data

MDS 3.0 QMs are publicly reported quarterly on Nursing Home Compare, as were the previously endorsed QMs based on the MDS 2.0. Unless otherwise specified, the analyses in this report are based on the episode file that RTI created for calculating the QMs for the fourth quarter of 2011. The episode file primarily comprises of MDS 3.0 assessments from July 1st to December 31st (Quarters 3 and 4 2011), but also includes some assessments from the second quarter of 2011 for some long-stay residents. Detailed definition of episode is presented in Section 2.4.3. The episode files used for each analysis are presented in the footnote of each table. Several analyses also used the episode files for calculation QMs for the previous quarters. For example, as a validity check we compared QM score across multiple quarters to examine seasonal variation. These analyses are based on the four quarters of data for calendar year 2011. This is sufficient data to evaluate most key factors except the potential for seasonal variation. Although MDS 3.0 data were collected starting in October 2010, analyses do not include that quarter for a few reasons. The data were expected to become more reliable as nursing home staff become familiar with the new and revised MDS 3.0 item set. Also, for the long-stay QMs, 3 months of data were insufficient to obtain a representative set of residents who had accumulated 100 days in facility.

1.3.3 Analytic Approach

In this section, we describe in detail the specific analytic approaches we used to test the new QMs, which largely follow the NQF guidance mentioned previously, adding a few considerations special to the MDS 3.0 measures.

Variability. For a QM to be effective, it should be able to distinguish between high-quality and low-quality nursing homes, and the QM should show sufficient variation to suggest

there is underlying variation in quality across providers. Using fourth quarter 2011 data, the variability in reported rates across nursing homes was analyzed, focusing on the interquartile range (the difference between the 75th and 25th percentiles) and the percentage of nursing homes with "perfect scores" (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes). These analyses help identify potential clustering of nursing homes across the distribution of the measure, including potential ceiling effects, if a large proportion of homes have perfect scores.

Reportability. Small facilities may not have enough cases to provide statistically large enough samples to calculate their QMs. CMS sets minimum sample sizes of 20 resident episodes for short-stay measures and 30 resident episodes for long-stay measures. We conducted analyses examining the proportion of nursing homes with sufficient sample sizes to meet minimum public reporting requirements, called "reportability" here, for each QM.

Reliability. Reliability refers to the consistency of results obtained when a measurement is repeated or the vulnerability of a measurement to random error. Dramatic changes in nursing home QM scores from one quarter to the next may indicate measure instability rather than true changes in nursing home quality. We compared results for each facility across multiple quarters of data. We evaluated the percentage of facilities showing large changes in facility ranking (three deciles or more) and in QM scores from quarter to quarter.

Validity. It is important that the QMs and individual items used to specify the QMs measure what they purport to measure, that is, the measures and items are valid. Several validity tests were conducted for each QM to examine its ability to measure true underlying variation in quality across facilities. To test the concurrent validity or the correlation of a measure with a QM measuring a similar aspect of nursing home quality (e.g., short-stay and long-stay pain QMs) of the measures, correlations among the paired and related measures were performed. Clinically related QMs are expected to vary similarly for any given facility. That is, a facility's percentile ranking for one QM would be correlated with another clinically related QM. These analyses required multiple quarters of data. To examine the potential that variation in the QMs could be attributed to other sources, we also examined the proportion of variance that could be attributed to the state in which the nursing home was located and the potential influence of the season/quarter on QM scores.

Risk Adjustment. For a QM to be valid and equitable, it should be designed to take into account sources of variation that are outside of a nursing home's control. For example, for outcome measures, it is important to consider the role of variation in resident characteristics in determining risk prior to comparing outcomes across different populations. Implementation of the MDS 3.0 provided new opportunities to develop and test new risk-adjustment models. Risk-adjustment testing included both: (1) denominator exclusions and (2) covariates. Analyses also investigate the logistic regression previously developed and changes using MDS 3.0 data. Finally, we also examined the impact of using different risk-adjustment calculation methods, including the classical logistic regression used for the MDS 2.0 QMs: fixed-effects and random-effects logistic models. Additional details can be found in the following sections specific to individual QMs.

Additional Analyses. This report also includes other analyses including:

MDS 2.0 QM Comparison. We compared facility MDS 3.0 QM results with their QM performance using MDS 2.0 data. For QMs whose underlying items did not change significantly from the MDS 2.0 to MDS 3.0 (e.g., physical restraints), it was expected that facilities would have similar results. For QMs which measured the same concept but whose underlying items changed significantly (e.g., pressure ulcers; Stage 1 pressure ulcers are no longer reported) we expected differences in the facility scores. Here we examined changes in national distributions in facility-level scores for MDS 2.0 and MDS 3.0 QMs which were comparable at both the item and construct level. These analyses are largely descriptive as differences in facility and national-level rates may be attributable to temporal trends in overall care and clinical practice or changes in patient population rather than differences in QM definition and difficult to disentangle where definitions varied greatly from MDS 2.0 to MDS 3.0.

Missing Data. We examined the extent, distribution, and impact of missing data on the measures to determine whether they affect facility-level measure scores. Specific analyses for the frequency of missing data examined: (1) type of assessment (with a focus on discharge assessments, which are new in MDS 3.0); (2) impact on QMs (i.e., analyze underlying items that make up the measures); (3) impact on QM specific exclusions; (4) facility-level reporting (e.g., do missing data exclude so many residents from the denominator that facility-level score cannot be reported?); (5) differences in missing data patterns between long-stay episodes compared with short-stay episodes; and (6) patterns of missing data on interview items for residents who cannot be understood, who have dementia, or who are cognitively impaired.

Short-Stay and Long-Stay Definitions. A significant change in the nursing home QMs specifications between MDS 2.0 and MDS 3.0 was redefining the QM resident populations. The MDS 3.0 short-stay and long-stay definitions replace the MDS 2.0 post-acute and chronic populations. In Section 2 we examine the facility-level distribution of long- and short- stay populations, as well as the distribution of key resident characteristics in the two populations to help inform evaluation of the appropriateness of these new resident population definitions.

Including Discharge Assessments in QM Specification. The MDS 3.0 includes a new discharge assessment not previously used. The discharge assessment, with an identical set of items and clinical sections as other assessment types, was added to better account for continuity of care and clinical status at discharge from the facility. Discharge assessment items are included in the MDS 3.0 QM definition. We examined the impact of including discharge assessments in QM construction on facility reportability. Theoretically, more short-stay residents should be eligible for inclusion in QM denominator construction with the availability of the discharge assessments. Many short-stay residents discharged after the completion of the 5-day assessment and before the 14-day assessment would have been excluded for QMs that use model-based risk adjustment because they have only one assessment (model-based risk adjustment uses values from the initial assessment, which must be different than the target assessment). By including the discharge assessment in the MDS 3.0 QM specifications, these same short-stay residents now have a minimum of two assessments available for calculating the QM. Additionally, the discharge assessment will ensure that the most current information regarding residents' clinical and functional status at the end of their nursing home stay is captured (as opposed to using the prior assessment), previously unavailable under MDS 2.0. For

example, the discharge assessment provides current clinical information on short-stay residents discharged before their 30-day PPS assessment, or long-stay residents discharged during the 3 months between their OBRA assessments. In the past, the most current (prior) assessment for these residents would represent clinical information that is 30 to 90 days old. We also investigated the impact of incomplete discharge assessment data related to facility QM percentile score (e.g., QMs calculated with and without discharge assessments) and facility reportability. For example, information may be incomplete until facilities become accustomed to filling out the new discharge assessment. Furthermore, unexpected discharges may result in missing data because certain assessment items were not possible to complete (e.g., interview items).

1.3.4 Report Organization

Quality measure analyses regarding the dimensions above are discussed in the next four sections in this report. Section 2 provides a description of the MDS 3.0 data set and the short-stay and long-stay resident populations. Section 3 presents analyses of the 16 NQF endorsed quality measures. Section 4 provides analytic details at the individual quality measure level. Lastly, Section 5 presents the conclusion discussion.

SECTION 2
DESCRIPTION OF DATA SET AND SHORT-STAY/LONG-STAY RESIDENT POPULATIONS

2.1 Introduction of Long-Stay and Short-Stay Definitions

The current publicly reported QMs are calculated for distinct resident populations in nursing homes: short-stay and long-stay (formerly post-acute care and chronic care) residents. The separation of nursing home residents into short-stay and long-stay categories was based on the premise that residents in the two populations have important differences in their clinical and service needs. For example, short-stay residents are more likely to be recovering from an acute illness or medical procedure, such as surgery. Therefore, for example, they are more likely to present with pain and less likely to present with cognitive disorders than would their long-stay counterparts. These two populations often require different types of care. Consequently, calculation of short-stay measures and long-stay measures are intended to assess nursing home quality for substantially different populations.

With the introduction of the MDS 3.0 assessment, the definition of resident episodes included in short-stay and long-stay QMs changed. The short-stay and long-stay definitions are now based on length of stay (LOS) in the facility, measured by cumulative days in facility (CDIF) within an episode. (Note that an episode may consist of more than one stay separated by periods of time outside the facility, such as a hospitalization). Short-stay residents are defined as those with CDIF less than or equal to 100 days as of the end of the target period. Long-stay residents are defined as those with CDIF greater than or equal to 101 days as of the end of the target period. These short-stay and long-stay definitions are used in all materials associated with the nursing home QMs, including the NQF submissions, the RAI MDS 3.0 Manual, the Quality Management Information System (QMIS) documentation, and the QM specifications in the CMS *MDS 3.0 Quality Measures User's Manual*. The definition change was motivated by concerns about the accuracy of the information about Medicare payer information (type of assessment being Prospective Payment System [PPS]) on MDS assessments, upon which the MDS 2.0 definitions depended. Additionally, changes in the definition would increase the number of residents identified as short-stay and included in the short-stay QMs, which was a challenge in the MDS 2.0 QM specifications, particularly in terms of risk-adjustment methodology. The decision to base QMs on length of stay was finalized by CMS after careful consideration regarding analyses of data and in-depth conversations with stakeholders.

2.2 History of Long-Stay and Short-Stay Definitions

The MDS 2.0 nursing home QM population definitions were primarily based on the presence of certain types of MDS assessments required for a Medicare Part A stay (typically post-acute debilitative services). Medicare requires that nursing homes must assess the clinical condition of beneficiaries by completing the MDS Prospective Payment System (PPS) assessment for each Medicare resident receiving Part A SNF-level care for reimbursement under the SNF PPS. The presence of MDS PPS assessments, therefore, is a proxy of Medicare payer information and also purpose of care (i.e., rehabilitation). Specifically, residents were classified as short-stay based on the presence of a 14-day PPS assessment in their episode of care. Residents included in MDS 2.0 long-stay QMs were identified by the presence of a quarterly or annual MDS assessment.

There are several issues related to identifying short-stay residents and the calculation of the short-stay QMs in the MDS 2.0:

- Many short-stay residents leave before the 14-day PPS MDS assessment.
- The time period between the 5-day and 14-day PPS MDS assessments is sometimes very brief (as short as 3 days), which affects the ability to assess change in some QMs.
- Individuals who are ultimately long-stay residents are included in these measures, as are those who were in the nursing facility prior to the qualifying hospital admission for Medicare coverage.
- Evidence suggests that using PPS MDS assessments may not accurately identify residents receiving Medicare post-acute care. For example, a study found that using MDS PPS assessments tends to result in over-reporting the percentage of residents receiving Medicare Part A SNF services (Iowa Foundation for Medical Care & University of Michigan Stepwise Systems CareTrack, 2005).

The new MDS 3.0 short-stay definition addresses some of these issues by:

- Adding the discharge assessment to capture outcomes for residents who leave <u>before</u> the 14-day PPS MDS assessment is required
- Adding the discharge assessment captures outcomes for residents who leave <u>between</u> the 14-day PPS MDS assessment and the first quarterly MDS assessment.
- Using length of stay to define short-stay and long-stay populations addresses the potential misclassification caused by using PPS MDS assessments and potentially increases the short-stay population size.
- Although none of the changes introduced in the MDS 3.0 definition address the issue of residents initially classified as short-stay (and included in short-stay QM calculations) who eventually become long-stay in the subsequent periods, Section 2.3.5 analyses indicate that these residents have only a minimal impact on QM scores.

2.3 Detailed Definitions of Long Stay and Short Stay

The following definitions of short- and long-stay yield mutually exclusive groups that include all nursing home admissions. The definitions use a practical, operational approach that is consistent with other Medicare and Medicaid policies.

2.3.1 Short-Stay Definition

The short-stay population includes resident episodes of care of 100 or fewer cumulative days in the nursing home (days are not required to be consecutive to be considered part of the same episode of care). This definition encompasses each individual stay in the nursing home regardless of whether residents have multiple short-stay episodes in any nursing home. This definition corresponds to the length of the Medicare Part A benefit. This population definition does not require that residents be discharged at the end of their episode to be considered short-stay. Residents who died prior to accumulating 100 days in facility are also included in the short-stay population.

2.3.2 Long-Stay Definition

The long-stay population includes resident episodes in nursing homes comprised of more than 100 cumulative days. This population definition is not dependent on the type of assessments included in the episode.

2.3.3 Other Key Definitions

We present the following definitions below as they are critical to understanding how residents are classified as short- or long-stay and which assessments are used in calculating QMs.

Target Period. The span of time that defines the QM reporting period (e.g., a calendar quarter).

Stay. The period of time between a resident's entry into a facility and either (a) a discharge, or (b) the end of the target period, whichever comes first. A stay is also defined as a set of contiguous days in a facility. (Note that the short-stay and long-stay definitions are based on episodes, which may include more than one stay, and therefore be comprised of non-contiguous days in facility).

Episode. A period of time spanning one or more stays. An episode begins with an admission and ends with either (a) a discharge, or (b) the end of the target period, whichever comes first.

Target Assessment. The latest (i.e. most recent) assessment contained within the resident's selected episode, that has a target date (assessment reference date, or discharge or entry date) that is before the end of the target period and also no more than 120 days before the end of the episode.

OBRA Assessment. OBRA assessments are federally mandated and completed on all residents in Medicare and/or Medicaid certified nursing homes.

Prospective Payment System (PPS) Assessment. PPS assessments are Medicare required. They provide clinical information on beneficiaries receiving Part A SNF-level care. The assessments are necessary in order to be reimbursed under the SNF PPS for both SNFs and Swing Bed providers.

Discharge Assessment. The discharge assessment is required for a resident who is discharged from the facility. This assessment includes clinical items for quality monitoring as well as discharge tracking information.

Cumulative Days in Facility (CDIF). CDIF is the total number of days within an episode during which the resident was in the facility. It is the sum of the number of days for all stays included in an episode. If an episode consists of more than one stay separated by periods of time outside the facility (e.g., hospitalizations), only those days within the facility would count towards CDIF. Any days outside of the facility (e.g., hospital, home, etc.) would not count towards the CDIF total.

2.4 Analyses of Short-Stay and Long-Stay Resident Populations

In this section, we describe findings from analyses of short-stay and long-stay populations based on MDS 3.0 data files used to create QMs for the fourth quarter of 2011. Analyses focus on the types of assessments that are identified most often as target assessments in the two populations. Additionally, we examine resident demographic characteristics and lengths of stay. The purpose of these analyses is to identify differences and similarities between the two populations and provide justification for using the current definitions.

2.4.1 Descriptive Statistics

We first examine all the episodes with target assessments used in the construction of the MDS 3.0 QMs in the fourth quarter of 2011. Out of roughly 2.6 million resident episodes, slightly more episodes are classified as short-stay than as long-stay (1.4 million, or 53.4 percent, and 1.2 million, or 46.6 percent, respectively),[2] although it is important to keep in mind that the short-stay population is based on episodes ending in both the target (Quarter 4, 2011) and prior quarters (Quarter 3, 2011) for sufficient sample size, whereas the long-stay population is based only on episodes ending in the target quarter (Quarter 4, 2011).

Table 2-2 presents the distribution of target assessments by types of MDS assessment and by population (short-stay vs. long-stay) in the Quarter 4, 2011, QM definition. The analysis has two goals: (1) to examine how often the discharge assessment serves as the target assessment and (2) to compare the assessment record type of the target assessments for the short-stay and long-stay populations. The results show four overarching themes:

- The discharge assessment serves as the target assessment for about three quarters of short-stay residents, but less than 10 percent of long-stay residents.
- More than 65 percent of the target assessments of short-stay residents are "Discharge, return not anticipated," compared to only about three percent of the long-stay resident target assessments. This contrast suggests a difference in health characteristics and care goals between the residents in the short-stay and long-stay populations.
- While the target assessment for the short-stay episodes is most frequently a discharge assessment, the target assessment for long-stay episodes is most frequently a quarterly OBRA assessment.
- A PPS assessment serves as the target assessment more often for short-stay residents (23.8 percent) than for long-stay residents (6. percent), suggesting that short-stay residents are more likely to be receiving Medicare post-acute care services.

[2] RTI analysis of Quarter 3, Quarter 4, 2011 MDS 3.0 data (quarter_4_5\db144_request.log)

Table 2-2
Proportion of target assessments by type of assessment variables by short-stay and long-stay definition by record type

Record type[a]	Short-stay *n*[b]	Proportion of total short-stay (%)	Long-stay *n*[b]	Proportion of total long-stay (%)
PPS	329,797	23.8	74,272	6.0
Quarterly	17,729	1.3	790,966	64.3
Comprehensive	14,937	1.1	255,432	20.8
Discharge, return not anticipated	904,711	65.2	41,853	3.4
Discharge, return anticipated	139,062	10.0	66,688	5.4
Total	1,381,028		1,203,726	

[a] Entry, OMRA, and death in facility records cannot be used as target assessments. Record type is generated based on the field ITM_SBST_CD and A0310F, see MDS 3.0 QM User's Manual (Appendix B) for more detail.

[b] Column *n*'s do NOT add up to total because one target assessment can belong to multiple categories, e.g., PPS combined with comprehensive.

Analysis date: 6/11/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011(quarter_4_5\db144_request\db144_request.log)

2.4.2 Relative Proportions of Short-Stay and Long-Stay Residents within Facilities

Because the QMs were designed to capture two distinct resident populations, it is of interest to investigate whether nursing homes tend to provide services to a mix of short-stay and long-stay residents or if facilities tend to specialize in providing services to one type of resident (i.e., short-stay or long-stay). Because several MDS 3.0 QMs are paired in terms of long-stay and short-stay versions (e.g., Short-stay and Long-stay Pneumonia Vaccination) correlation analyses among these measures provide a validity check (detailed in Section 4). Facility specialization, focusing on care for short-stay or long-stay types of residents, could potentially be associated with weak (or even negative) correlation between short-stay and long-stay versions of the QM.

Table 2-3 shows descriptive statistics for the proportions of short-stay residents within nursing homes across MDS 3.0 Quarter 3 and 4, 2011, data. The facility-level percentile distribution of short-stay residents illustrates that some facilities provide services largely to a short-stay population (77.5 percent are short-stay residents in facilities at the 90th percentile), some facilities provide services largely to a long-stay population, but most provide services to a mix of short-stay and long-stay residents. The median facility proportion of short-stay residents is 42.5 percent; the mean is 45.5 percent. Given that slightly more than half of all episodes are short-stay (53.4 percent, ***Table 2-1***), it is likely that there are some larger or higher volume facilities that primarily care for short-stay residents. We analyze this by plotting facilities' proportion of short-stay residents against facility size. ***Figure 2-1*** shows that as facility size increases the proportion of short-stay residents increase slightly. In addition, very small facilities (50 beds or less) tend to have a higher proportion of short-stay residents compared to all other size facilities.

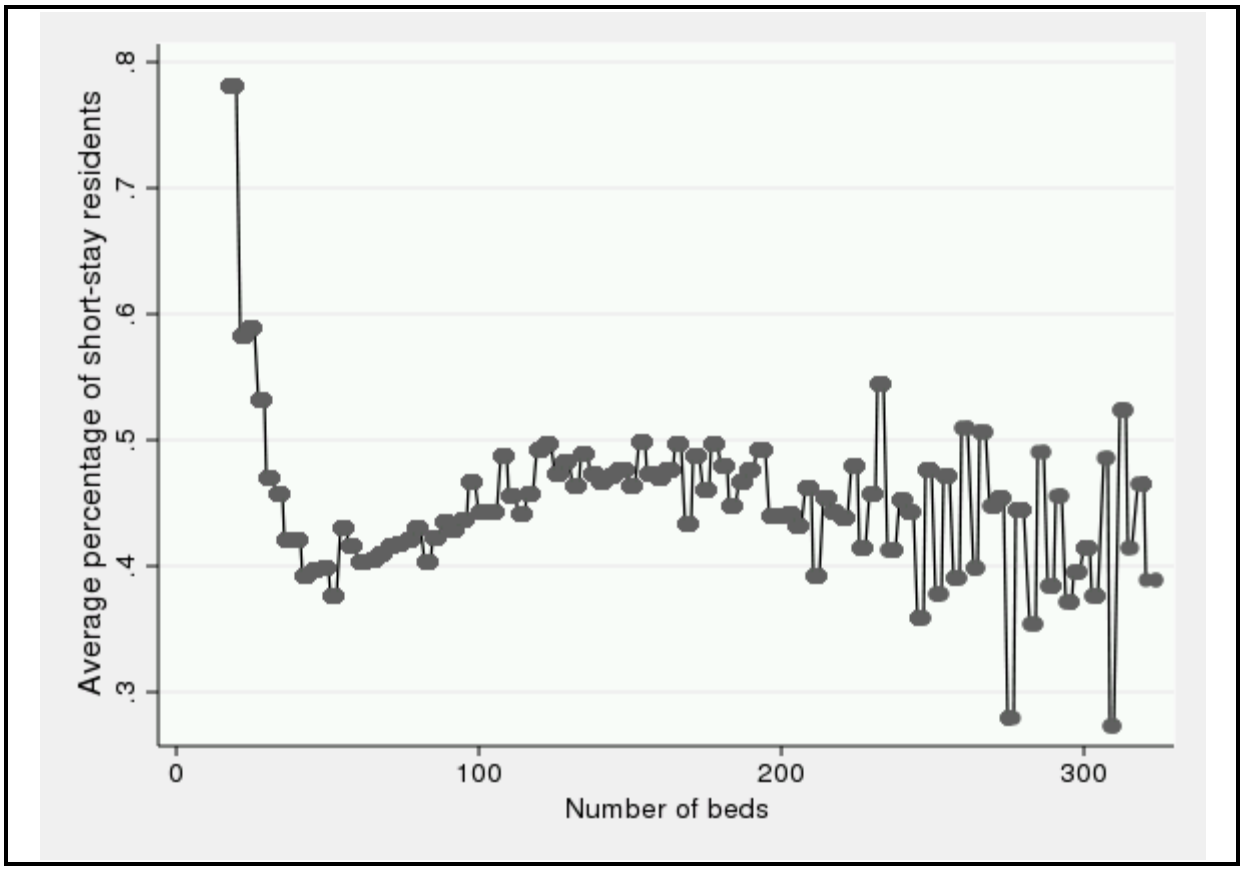
Table 2-3
Facility-level proportion of short-stay residents, fourth quarter of 2011

n (facilities)	Mean	SD	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,686	45.5%	22.4%	18.5%	28.8%	42.5%	59.8%	77.5%

Analysis date: 6/11/2012

SOURCE: RTI analysis of Quarter 3, Quarter 4, 2011 MDS 3.0 data episode file for Quarter 4, 2011 (quarter_4_5\db144_request\db144_request.log)

Figure 2-1
Facility-level proportion of short-stay residents by facility volume



Analysis date: 8/10/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db132_request\graph_db132.png)

2.4.3 Distribution of Length of Stay and Resident Characteristics by Long Stay and Short Stay

In this section we investigate whether the MDS 3.0 definitions of short-stay and long-stay differentiate residents with respect to health conditions and care goals. We compared the long-stay and short-stay populations by: (1) length of stay; (2) resident expectation of care; and (3)

health characteristics. The length of stay is calculated as CDIF (see Section 2.3). Resident health characteristics are extracted from the baseline of the target period (i.e., prior assessment for the long-stay residents and initial assessment for short-stay residents). The residents' overall expectation of care is based on item Q0300A.

Length of Stay. ***Table 2-4*** shows the distributions of CDIF by short-stay and long-stay residents included in the QM definition for Quarter 4 of 2011.

Table 2-4
Length-of-stay (cumulative number of days in facility) distributions for short-stay and long-stay resident episodes, fourth quarter of 2011

Length of stay	Target assessments (*n*)	Mean	SD	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
Short-stay	1,381,028	30.1	23.4	7	13	22	42	67
Long-stay	1,203,726	433.3	237.2	169	323	485	521	545

Analysis date: 6/11/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db145_request\db145_request.log; quarter_4_5\db146_request\db146_request.log)

Expectation of Care. To examine if MDS 3.0 definition of short-stay and long-stay differentiate residents with respect to their overall expectation of care (i.e., rehabilitative vs. long-term services and supports [LTSS]), we calculated the percentage of residents with a PPS assessment during the target period. A PPS assessment is required by Medicare for services reimbursable under Part A and therefore can be an indicator of receiving rehabilitative services. The vast majority of short-stay residents (81.4 percent) have a PPS assessment, but only 6.0 percent of long-stay residents do.[3] This small percentage of long-stay residents may have been discharged to an acute hospital and returned as anticipated. This distinction suggests that short-stay residents are mainly under rehabilitative care, while most long-stay residents receive LTSS care.

Sociodemographic and Health Characteristics. To ascertain differences in the short-stay/long-stay populations, we also compared health characteristics between the two populations. ***Table 2-5*** depicts the distribution of selected resident characteristics in the short-stay and long-stay populations. Although a more comprehensive list of characteristics were examined, we only present characteristics that (1) show substantial difference between the two samples and/or (2) are potentially related to differential numerator triggering rates for one or more QMs. For example, cognitive impairment and cancer may be associated with differential likelihood of having pain. Some health characteristics examined but are not presented include anemia (29.5 percent for long-stay vs. 28.7 percent for short-stay), heart failure (20.2 percent vs. 19.9 percent), thyroid disorder (20.1 percent vs. 18.8 percent), gastroesophageal reflux disease

[3] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 \quarter_4_5\db148_request\db148_request.log

Table 2-5
Resident characteristics by short-stay and long-stay

Resident characteristic	Short-stay *n* = 1,381,028 proportion	Long-stay *n* = 1,203,726 proportion
Age		
Age < 65	14.7%	15.2%
Age 65-74	20.0%	13.8%
Age 75-79	14.8%	11.1%
Age 80-84	19.0%	16.7%
Age 85+	31.6%	43.2%
Gender: Male	37.2%	31.5%
Race		
American Indian or Alaska Native	0.4%	0.4%
Asian	1.4%	1.7%
Black	10.2%	14.0%
Hispanic	4.1%	4.7%
Hawaiian/Pacific Islander	0.3%	0.3%
White	82.0%	77.0%
Cognitive Impairment*	14.6%	40.4%
Bed mobility, poor self-performance (Extensive or total dependence)	67.6%	62.0%
Transfer, poor self-performance (Extensive or total dependence)	68.4%	62.8%
Malnutrition or at risk of malnutrition	3.3%	2.3%
Specific Diagnoses		
Cancer	9.6%	4.5%
Coronary Artery Disease	26.1%	17.8%
Peripheral Vascular Disease or Peripheral Arterial Disease	7.7%	8.8%
Renal Insufficiency, Renal Failure, or End-Stage Renal Disease	14.7%	8.5%
Pneumonia	9.1%	2.8%
Urinary Tract Infection	16.5%	8.1%

(continued)

Table 2-5 (continued)
Resident characteristics by short-stay and long-stay

Resident characteristic	Short-stay *n* = 1,381,028 proportion	Long-stay *n* = 1,203,726 proportion
Diabetes Mellitus	33.5%	32.4%
Arthritis	28.2%	27.3%
Osteoporosis	12.7%	17.8%
Hip Fracture	7.9%	2.1%
Other Fracture	10.0%	2.5%
Alzheimer's Disease	4.4%	17.7%
Cerebrovascular Accidents, Transient Ischemic Attack, or Stroke	12.0%	17.9%
Dementia	17.8%	45.8%
Anxiety Disorder	17.2%	23.9%
Depression	29.8%	51.1%
Manic Depression	2.1%	4.3%
Psychotic Disorder	2.6%	11.0%
Schizophrenia	1.6%	7.1%
Asthma, COPD, or Chronic Lung Disease	23.6%	19.8%

* A resident is cognitively impaired if s/he has BIMS ≤ 7 or has short-term memory problem AND severely impaired cognitive skills for daily decision making.

Analysis date: 6/12/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db148_request\db148_request.log)

or ulcer (30.4 percent vs. 28.0 percent), cerebral palsy (0.2 percent vs. 1.0 percent), respiratory failure (2.6 percent vs. 1.6 percent) and multiple sclerosis (0.6 percent vs. 1.4 percent). Results indicate short-stay residents differ from long-stay residents across a wide array of resident characteristics (MDS 3.0 items). Short-stay residents tend to be younger, more often male and white. Short-stay and long-stay residents have different health profiles. For example, short-stay residents exhibit higher rates of intact cognitive status and are slightly less independent in physical function. Certain health conditions are more prevalent among short-stay residents, such as cancer, coronary artery disease, and pneumonia. Health conditions most prevalent among long-stay residents include Alzheimer's disease and dementia. The difference in the prevalence of these cognitive disorders is substantial between long-stay and short-stay residents.

2.5 Short-Stay Population Analyses

A persistent concern regarding the QM short-stay and long-stay definition is that, even with the MDS 3.0 definition refinements, the focus on length of stay could still result in misclassification of true long-stay residents (i.e., residents receiving LTSS care) as short-stay

residents. For example, long-stay residents in their first 100 days of nursing home stay are included in the short-stay definition when calculating short-stay QMs even if they eventually become long-stay residents in subsequent target period. In addition, residents receiving LTSS care who died before accumulating 100 cumulative days in facility are included in the short-stay QM definition.

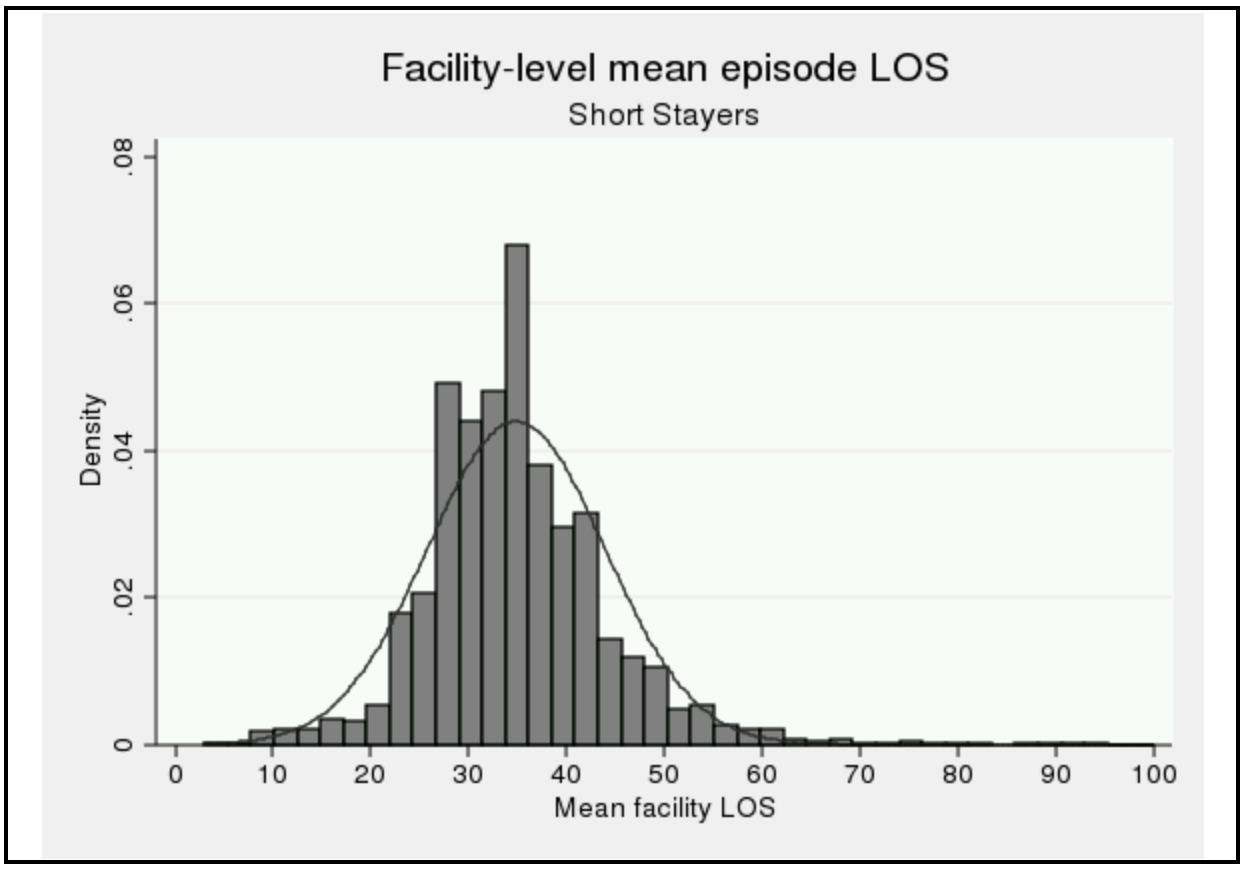
We next conducted a variety of analyses within the short-stay population to gauge the extent of potential LOS misclassification. We first stratified the short-stay population by whether residents had their episode initiated with a PPS assessment. Although we recognize the previously mentioned issue with the validity of using the presence of a PPS assessment to identify a resident admitted to a nursing home after a prior acute inpatient hospital stay, it is likely the set of residents with no PPS assessments would be relatively unlikely to be receiving services in a nursing after an acute hospitalization. Short-stay residents without a PPS assessment initiating their episode may be LTSS residents in their first 100 days of nursing home stay and/or later die before accumulating 100. We compared resident characteristics between the two subgroups of short-stay residents for evidence that residents with and without initial PPS assessments differ on key characteristics, and if residents without initial PPS were more similar to long-stay residents. Next we looked at the differences in resident expectations for care, among the short-stay residents, stratified by whether residents had a PPS assessment or not as their initial assessment in their episode, and in comparison with the long-stay population. If misclassification is an issue with the current resident population definitions, we would expect that resident expectations for care among short-stay residents with no PPS assessment to be more similar to long-stay residents. We also looked at the proportion of residents in the short-stay population who died before accumulating 100 days in facility, because these residents may be more similar to long-stay residents than "true" short-stay residents. Additionally, we looked at mean length of stay by initial assessment record type. We assumed that short-stay residents with no PPS assessment, if they were more similar to long-stay residents, should have systematically longer lengths of stay than short-stay residents with a PPS assessment. Lastly, we identified the set of residents in the short-stay population who were reclassified as long stay in subsequent quarters and examined whether including these residents in the short-stay population definition introduces bias into the short-stay QMs.

2.5.1 Distribution of Length of Stay and Resident Characteristics—Short-Stay Analyses

We examined LOS by short-stay residents subsets to ascertain the magnitude of short-stay residents who eventually go on to become long-stay residents but who are initially included in short-stay QM calculations. We also examined the distribution of facility-level mean CDIF for the short-stay sample to ascertain if some facilities have disproportionately more short-stay residents with LOS close to 100 days (that is, a disproportionate share of short-stay residents who are most likely to transition to long-stay residents). For the short-stay population, data from ***Table 2-4*** show the difference between the mean (30.1 days) and the median (22 days) indicates a positively skewed (longer LOS) distribution suggesting the short-stay population does include some residents who eventually accumulate 100 days in facility and qualify as long-stay. However, ***Figure 2-2*** (which shows the distribution of facility-level mean CDIF for the short-stay residents) does not suggest a bimodal distribution (using a probability density function—a function that describes the relative likelihood for the variable to take on any given value in the

distribution), which would support that fact that a substantial proportion of facilities have a large number of residents who will eventually qualify as long stay. The distribution is largely normal (mean of 35 days; median of 34 days) with a slight negative skew and a long right-side tail suggesting a small number of facilities may be more likely have short-stay residents who eventually go on to become long-stay residents compared with other facilities. Section 2.5.3 provides additional analyses illustrating that when short-stay residents are tracked over numerous quarters, 10.6 percent of residents classified as short-stay at the end of Quarter 3, 2011, accumulated sufficient days in facilities during Quarter 4, 2011, to be reclassified as long stay in subsequent target periods (and thus, would be captured in the long-stay QMs)

Figure 2-2
Discrete probability density of cumulative number of days in facility: short-stay resident episodes



Analysis date: 6/11/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db145_request\db145_request.png)

Within the short-stay population, residents who do not start their episode with a PPS assessment may in fact receive LTSS care and have health characteristics different from rehabilitative short-stay residents. We, therefore, compared length of stay and resident characteristics between short-stay residents whose initial assessment is PPS and those whose initial assessment is non-PPS.

Short-stay residents with and without a PPS assessment have very similar length of stay (30.3 days for residents without a PPS assessment, and 30.5 days for residents with a PPS assessment).[4] In addition, ***Table 2-6*** shows that the two groups of short-stay residents are largely similar with regard to health characteristics except for a few conditions. For example, compared with short-stay residents without a PPS assessment, those with a PPS assessment were slightly more dependent in physical function and have higher prevalence of coronary artery disease, hypertension, pneumonia, urinary tract infection, and arthritis. However, the distinction between the two populations is small and does not demonstrate the clear long-stay/short-stay health condition stratification illustrated in the previous section. These findings suggest that the short-stay residents whose initial assessment is non-PPS are more appropriate to be included in the short-stay population than in the long-stay population.

Table 2-6
Short-stay resident characteristics by stratified by type of the initial assessment

Resident characteristic	Initial assessment = PPS *n* = 801,720 proportion	Initial assessment = non-PPS *n* = 183,341 proportion
Age		
Age < 65	10.7%	31.8%
Age 65-74	20.8%	16.6%
Age 75-79	15.4%	11.9%
Age 80-84	19.0%	15.0%
Age > 84	33.1%	24.7%
Gender: Male	36.5	40.3%
Race		
American Indian or Alaska Native	0.3%	0.6%
Asian	1.2%	2.1%
Black	9.4%	13.7%
Hispanic	3.4%	6.8%
Hawaiian/Pacific Islander	0.2%	0.4%
White	83.7%	74.3%
Cognitive impairment*	14.5%	14.8%
Bed mobility, poor self-performance (Extensive or total dependence)	68.7%	62.8%
Transfer, poor self-performance (Extensive or total dependence)	69.5%	63.8%
Malnutrition or at risk of malnutrition	3.3%	3.0%

(continued)

[4] SOURCE: RTI analysis of MDS 3.0 data.\lkomp\db requests\db179_01.xls

Table 2-6 (continued)
Short-stay resident characteristics by stratified by whether initial assessment in episode is PPS or non-PPS

Resident characteristic	Initial assessment = PPS *n* = 801,720 proportion	Initial assessment = Non-PPS *n* = 183,341 proportion
Diagnoses		
Cancer	9.7%	9.1%
Coronary Artery Disease	27.4%	21.3%
Hypertension	75.1%	70.8%
Peripheral Vascular Disease or Peripheral Arterial Disease	7.9%	7.0%
Renal Insufficiency, Renal Failure, or End-Stage Renal Disease	15.4%	12.3%
Pneumonia	9.6%	6.7%
Urinary Tract Infection	17.2%	13.2%
Diabetes Mellitus	33.3%	34.6%
Arthritis	28.9%	25.5%
Osteoporosis	13.2%	11.1%
Hip Fracture	8.3%	6.1%
Other Fracture	9.9%	10.0%
Alzheimer's Disease	4.3%	4.7%
Cerebrovascular Accident, Transient Ischemic Attack, or Stroke	12.0%	12.3%
Dementia	18.0%	16.7%
Anxiety Disorder	17.1%	17.5%
Depression	29.5%	30.9%
Manic Depression	1.9%	3.0%
Psychotic Disorder	2.5%	3.1%
Schizophrenia	1.4%	2.7%
Asthma, COPD, or Chronic Lung Disease	24.0%	22.1%

* A resident is cognitively impaired if s/he has BIMS ≤ 7 or has short-term memory problem AND severely impaired cognitive skills for daily decision making.

Analysis date: 6/12/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db148_request\db148_request.log)

We compared the distribution of resident's' overall expectations/goals of care (as measured by MDS 3.0 item Q0300A, e.g., discharged to the community) by long-stay and two short-stay populations. This comparison begins to explore whether MDS 3.0 definitions of short-stay and long-stay also reflect "intent to treat" (whether a resident receives post-acute/rehabilitative care or long-term/custodial care), thus differentiating between residents with

distinctly different care goals. ***Table 2-7*** illustrates that a large majority of short-stay residents expect to be discharged to the community, while a majority of the long-stay residents expect to remain in the facility. The small percentage of short-stay residents (13.6 percent) expecting to remain in the facility could reflect the 10.6 percent of residents classified as short-stay at the end of Quarter 3 2011 who accumulated sufficient days in facility during Quarter 4, 2011, to be re-classified as long-stay in the subsequent period (presented in 2.5.4). The difference between the short-stay and long-stay residents' overall expectation of care indicates that the MDS 3.0 definition of short-stay and long-stay identify nursing home residents with different care goals. We also found that expectations of care for short-stay residents without a PPS assessment are more similar to short-stay residents with a PPS assessment than to long-stay residents. This finding supports the use of length of stay to define short-stay and long-stay residents.

Table 2-7
Distribution of resident's overall expectation of care by short and long stay

Resident's overall expectation	Short stay, all *n* (%)	Short stay, PPS *n* (%)	Short stay, non-PPS *n* (%)	Long stay *n* (%)
Expects to be discharged to the community	1,099,283 (75.8%)	827,836 (79.1%)	271,447 (67.2%)	281,523 (18.4%)
Expects to remain in this facility	160,752 (11.1%)	83,296 (8.0%)	77,456 (19.2%)	1,059,481 (69.4%)
Expects to be discharged to another facility/institution	47,391 (3.3%)	34,695 (3.3%)	12,696 (3.1%)	24,135 (1.6%)
Unknown or uncertain	107,722 (7.4%)	78,182 (7.5%)	29,540 (7.3%)	131,277 (8.6%)
Missing	34,786 (2.4%)	22,084 (2.1%)	12,702 (3.2%)	31,032 (2.0%)
Total	1,449,934	1,046,093	403,841	1,527,448

NOTE: Residents may have multiple assessments with overall expectation data.

Analysis date:7/9/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db140_request\db140_request.log)

2.5.2 Short-Stay Residents Who Die Before Accumulating 100 CDIF

Residents who die before accumulating 100 CDIF may be frail individuals at the end of life who were admitted for long-term services and support in the nursing homes but misclassified into the short-stay category. We therefore report the facility distribution of percent of short-stay residents who die before accumulating 100 CDIF in ***Table 2-8***. The average facility percent of short-stay residents who die before accumulating 100 CDIF is about 6.7 percent; the facility median is 5.3 percent. The percentage varies across facilities, suggested by a standard deviation of 7.1 percent and a interquartile range of 7.7 percent. The across-facility variation in the percent of short-stay residents who die before accumulating 100 CDIF suggest that some facilities may have more residents with long-term care needs misclassified into the short-stay category, although the overall percentage is low (10-90 percentile = 0-15 percent). It is problematic, however, to assume that all residents who died during a short-stay were residents with needs more similar to those in a long-stay population (i.e., long-term services and support). Residents

admitted after acute hospital stays are likely to have an elevated risk for mortality as well. To further investigate, we also compared the percentage of residents who died before accumulating 100 days in facility for short-stay residents with and without an initial PPS assessment. About 4 percent of short-stay residents with a PPS assessment die before accumulating 100 CDIF, compared with 3.4 percent of short-stay residents without a PPS assessment.[5]

Table 2-8
Facility distribution of percent of short-stay residents who die before accumulating 100 cumulative days in facility

—	*n* (facilities)	Mean	SD	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
All	15,636	8.8%	7.8%	0.3%	3.6%	7.2%	12.1%	18.1%

Analysis date: 7/24/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db147_request\new_run\db147_request_v1.log)

2.5.3 Early Long-Stay Residents

In this section we examine the impact on the short-stay QMs of including residents who qualify as short-stay during an initial quarter who subsequently become long-stay (hereafter referred to as "early long-stay"). There is some concern that these early long-stay residents may be clinically different than the 'true' short-stay residents thus, potentially biasing the calculation of the short-stay QMs. Based on Quarter 3 and 4, 2011 data we subset residents who were classified as short-stay residents in Quarter 3 2011 who were subsequently classified as long-stay residents in Quarter 4 of 2011. That is, these residents, though classified as short-stay in the Quarter 3 2011 calculations of the QMs, were actually long-stay residents within the first 100 days of their stay. We identified 145,592 long-stay residents who were in the first 100 days of their stay during Quarter 3 of 2011, representing 10.6 percent of the short-stay population counted in that quarter. We used this population of "early long-stay" residents to further explore mean facility scores for the short-stay QMs.

We have included analyses of four short-stay QMs below. ***Table 2-9*** illustrates the proportion of early long-stay residents in the numerator and denominator for each short-stay QM. We then compare resident-level numerator triggering rate (the percentage of residents in the denominator who also counted in the numerator for each QM) and mean facility scores before and after excluding "early long-stay" residents from each QM. Table 2-9 (2nd and 3rd columns from left) also illustrates a slightly higher percentage of early long-stay residents included in the numerator (20.0 percent) and the denominator (12.3 percent) for short-stay pain. Overall, excluding 'early long-stay' residents from the QM calculation results in very little difference in the facility QM scores for all four short-stay measure. The change in facility-level QM score mean for each of these measures is less than one percentage point, and the mean does not change at all for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay). Thus, the presence of early long-stay residents appears not to have a substantial effect on scores for short-stay quality measures.

[5] SOURCE: RTI analysis of Quarter 3 and Quarter 4, 2011, MDS 3.0 Data (quarter_4_5\db147_request\new_run\db147_request_v1.log)

Table 2-9
Impact of excluding early long-stay residents on short-stay quality measure scores, time period

Quality measure name	Early long-stay residents: numerator *n* (%)	Early long-stay residents: denominator *n* (%)	Resident-level numerator triggering rate: based on all short-stay residents with CDIF ≤ 100 %	Resident-level numerator triggering rate: excluding early long-stay residents %	Facility-level quality measure mean score: based on all short-stay residents with CDIF ≤ 100 % (SD %)	Facility-level quality measure mean score: excluding "early long-stay" residents % (SD %)
QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Total denominator = 973,812)	23,851 (20.0%)	119,556 (12.3%)	23.7%	24.2%	22.9% (12.3%)	23.5% (13.1%)
QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) (Total denominator = 1,322,244)	2,783 (12.4%)	138,096 (10.4%)	1.7%	1.7%	1.9% (2.4%)	2.0% (2.7%)
QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) (Total denominator = 828,843)	51,844 (8.1%)	75,586 (9.1%)	77.2%	78.0%	76.5% (24.5%)	77.2% (24.3%)
QM#0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay) (Total denominator = 1,377,201)	112,947 (9.9%)	145,592 (10.6%)	82.6%	83.2%	79.7% (22.0%)	79.7% (22.3%)

Analysis date: 6/12/2012 – 6/13/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db149_request\db149_request.log; quarter_4_5\db150_request\db150_request.log; quarter_4_5\db152_request\db152_request.log; quarter_4_5\db153_request\db153_request.log)

2.6 Summary and Conclusions

The MDS 3.0 short-stay and long-stay definitions were developed to categorize residents with important differences in clinical and service needs and to design QMs capable of assessing substantially different facets of nursing home quality. With the introduction of the MDS 3.0 long-stay/short-stay definition, resident categories in the QMs changed from one based on payer status to one based on length of stay in the facility. Short-stay residents are now defined as those with a length of stay, measured by cumulative days in facility, of 100 days or fewer; long-stay residents are defined as having a length of stay greater than 100 days. The 100 day cut-off was selected to align with the Medicare Part A benefit coverage—the major source of reimbursement for short-stays. The new definitions also address some criticisms of the MDS 2.0 definitions of the post-acute and long-term care categories. For example, MDS 3.0 definition of the short-stay includes outcomes for those residents who leave before the 14 day PPS assessment is required and outcomes for those who leave between the 14-day PPS assessment and the first quarterly assessment. Finally, the new definition, along with the introduction of the discharge assessment, allows for a potentially larger short-stay population in calculating the QMs.

In this section we reported on the distribution of type of target assessment, length of stay and resident characteristics for the short-stay and long-stay populations. The results indicate that short-stay and long-stay residents are different in many ways. Short-stay residents are more likely to have a discharge assessment or a PPS assessment as their target assessment, whereas long-stay residents more often have a quarterly or a comprehensive assessment as the target assessment. The average length of stay is about 30 days for short-stay residents and about 433 days for long-stay residents. Short-stay and long-stay residents have somewhat different demographic and health characteristics as well as overall expectations for their continued stay in the nursing home. The large majority of short-stay residents are expected to be discharged to the community, whereas the large majority of long-stay residents are expected to remain in the facility. These descriptive analyses suggest some differences between the MDS 3.0 definitions of short-stay and long-stay residents by age, gender, race, cognitive impairment, mobility, transfer, some diagnoses and resident's expectations of care.

We also analyzed the possibility of population misclassification. We examined (1) facility distribution of percent of short-stay residents who die before accumulating 100 cumulative days in facility, (2) length of stay by initial MDS assessment type and the impact on short-stay QM scores of excluding short-stay residents who eventually accumulate more than 100 days in facility, and (3) short stay residents who subsequently become long-stay residents. We find that some facilities may have more residents with long-term care needs (identified by non-PPS assessment type) misclassified into the short-stay population because these potential long-term care residents die before accumulating 100 cumulative days in facility. However, overall the facility mean percentage is low (8.8 percent). We also find that approximately 11 percent of residents classified as short-stay in an initial quarter go on to be reclassified as long-stay in the subsequent quarter. However, excluding these residents from calculating short-stay QM scores has negligible effects on all four short-stay QM scores.

SECTION 3
ANALYSES OF 16 NQF-ENDORSED MEASURES

In this section of the report, we examine several key aspects of the 16 NQF-endorsed MDS 3.0 QMs. These include the variability of the measures across facilities, reportability, the threat to validity posed by missing data, initial analyses of the reliability of the measures, and an examination of the impact of including discharge assessments in the set of assessments eligible to be used in the calculation of the QM scores.

3.1 Variability

A well-designed QM should be sensitive enough to capture the actual variability in quality of care among nursing facilities by discriminating between facilities with varying performance levels. If facilities' QM scores are found to cluster in a narrow range or at the top or bottom of the range of possible scores, true differences in performance may be either obscured or overstated. We examined the variability in reported MDS 3.0 QM rates across facilities. To measure the spread of the distribution of scores, we analyzed the interquartile range for the scores on each measure. A narrow interquartile range (the difference between 25th and 75th percentiles) suggests clustering of facilities on a given QM. It may, therefore, affect reporting of facility ranking on a QM, since a small change in QM score from one quarter to the next may result in an exaggerated shift in rank that exaggerates the change in how the facility performed. To examine possible ceiling effects, we examined the proportion of facilities with perfect scores (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes). ***Table 3-1*** shows the results of this analysis using the third and fourth quarters of 2011 MDS 3.0 data. For each of the 16 NQF-endorsed QMs, the table shows the mean score; the standard deviation; the 10th, 25th, 50th, 75th, and 90th percentiles; the interquartile range; and the percentage of facilities with perfect scores.

Key findings from these analyses include:

- Variability of QM scores, as measured by interquartile range, is highest for:
  - QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) (24.8 percent)
  - QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) (23.3 percent)
  - QM #0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay) (23.0 percent)
- Variability of QM scores is particularly narrow for measures with relatively low prevalence. The six measures with the lowest mean scores also had the six narrowest interquartile ranges. These are: QM #0678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay) (mean: 1.9 percent; interquartile range: 2.8 percent); QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) (mean: 2.4 percent; interquartile range: 3.2 percent); QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) (mean: 3.4 percent; interquartile range: 3.3 percent); QM #0686 Percent of Residents Who

Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) (mean: 4.1 percent; interquartile range 4.1 percent); QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) (mean: 7.1 percent; interquartile range: 5.8 percent), and QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) (mean: 6.9 percent; interquartile range: 6.2 percent).

- The highest proportions of perfect scores were for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) (50.5 percent); QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) (39.4 percent); and QM #0678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay) (29.3percent). Not surprisingly, these measures also show narrow interquartile ranges (3.2 percent, 6.4 percent, and 2.9 percent, respectively).
- QM scores were relatively high for the four vaccination measures, QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) (93.8 percent); QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) (89.1 percent; QM #0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay) (81.0 percent); and QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) (79.7 percent). Note that these four QMs, unlike the other QMs, are positive measures, for which higher scores indicate better quality.
- Fewer than 2 percent of facilities report perfect scores on the following measures: QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) (0.3 percent); QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Have Increased (Long Stay) (1.0 percent); and QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) (1.4 percent). These measures also demonstrate relatively wide interquartile ranges: 24.8 percent, 11.8 percent, and 17.2 percent, respectively.

Table 3-1
Quality measure score distribution by QM

Measure name	*n*	Mean score, %	SD	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range %
QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)	13,868	3.4	2.7	0	1.5	2.9	4.8	6.8	14.4	3.3
QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)	11,033	23.1	12.6	7.4	13.8	22.2	31.0	39.5	1.4	17.2
QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)+	11,896	11.5	8.5	1.9	5.0	9.9	16.5	23.2	4.4	11.5
QM #0678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay)+	10,763	1.9	2.4	0	0	1.3	2.9	4.9	29.3	2.9
QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)	11,352	6.9	4.7	1.7	3.3	6.3	9.5	12.9	6.9	6.2
QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay)	11,833	79.7	19.2	52.2	70.7	85.7	94	98.4	7.3	23.3
QM #0681 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Long Stay)	13,795	89.1	13.7	70.6	84	94.4	98.8	100	21.2	14.8
QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)	12,762	81.0	21.0	50	73.2	88.6	96.2	99.7	9.9	23.0
QM #0683 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)	13,870	93.8	12.4	82.5	93.6	98.4	100	100	39.4	6.4

(continued)

Table 3-1 (continued)
Quality measure score distribution by QM

Measure title	*n*	Mean score, %	SD	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range %
QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay)	13,773	7.8	5.7	1.3	3.5	6.7	10.8	15.2	7.2	7.3
QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)	6,689	42.8	17.3	20.3	30.4	42.6	55.2	65.5	0.3	24.8
QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)+	13,527	4.1	3.5	0	1.8	3.5	5.9	8.5	13.1	4.1
QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)	13,864	2.4	4.2	0	0	0	3.2	7.1	50.5	3.2
QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Have Increased (Long Stay)	12,735	16.7	9.3	6.3	10	15.4	21.8	29.0	1.0	11.8
QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)	13,761	7.1	4.6	1.8	3.8	6.5	9.6	13.0	5.7	5.8
QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)	13,775	7.3	10.8	0	0.7	3.8	9.5	18.2	24.3	8.8

+Model-adjusted risk adjustment applied. See Section 3.6 for discussion of risk adjustment methodology.

Analysis date: 4.17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (qm_quarter_4_5\complete)

3.2 Reportability for Each QM and Impact of Discharge Assessment on Reportability

In this subsection, we summarize the reportability for each QM and the impact of including discharge assessments on reportability. The discharge assessment is a new feature to the third generation of the MDS. Completing the MDS discharge assessment upon a resident's departure from a facility contributes another data point for QM analysis. If discharge assessments are properly filled out, they can provide assessment information in cases where no other assessment is available and thus the resident can be included in QM calculations. This not only results in a more accurate QM accounting, but also increases the likelihood that the facility would meet the sample size requirements for QM reporting. The completeness of the discharge assessment is paramount in meeting both these goals. Missing data is of particular concern in discharge assessments since residents may be discharged unexpectedly. Missing data in discharge assessments is further acerbated by nursing homes who may not have incorporated the discharge assessment into their formal assessment process (i.e., the new discharge assessment currently is not federally mandated nor associated with payment). To assess the impact of discharge assessments on reportability, we analyzed the impact of removing the discharge assessment from the calculation of the measures. ***Table 3-2*** presents the results of this analysis. For all measures but one, removing the discharge assessments from analyses led to a minimal reduction of the number of facilities able to report: for 14 of the 16 measures, this change represented 2 percent or fewer of the total number of facilities. Reportability actually increased for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) by 5.3 percentage points, owing to a high rate of missing items (and thus exclusions) on pain-related items on discharge assessments. However, the loss in reportability after excluding discharge assessments are substantial for QM #678, Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay). Discharge assessments helped 15.6 percent of facilities report that could not without the discharge assessment.

Table 3-2
Number and percent of nursing homes able to report quality measures with and without discharge assessments included in quality measure calculations

Measure name	With discharge assessments included *n* (Percent)	With discharge assessments excluded *n* (Percent)
QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)**	13,868 (88.4%)	13,860 (90.0%)
QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)	11,033 (70.3%)	11,851 (75.6%)
QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)	11,896 (75.8%)	11,879 (75.7%)
QM #0678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay)	10,763 (68.6%)	8,311 (53.0%)
QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)	11,352 (72.4%)	11,283 (71.9%)

(continued)

Table 3-2 (continued)
Number and percent of nursing homes able to report quality measures with and without discharge assessments included in quality measure calculations

Measure name	Discharge assessments included *n* (Percent)	Discharge assessments excluded *n* (Percent)
QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay)	11,833 (75.4%)	11,522 (73.5%)
QM #0681 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Long Stay)	13,795 (87.9%)	13,779 (87.8%)
QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)	12,762 (81.0%)	12,578 (80.2%)
QM #0683 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)	13,870 (88.4%)	13,862 (88.4%)
QM #0684 Percent of Long-stay Residents with a Urinary Tract Infection	13,773 (87.8%)	13,758 (87.5%)
QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)	6,689 (42.6%)	6,545 (41.7%)
QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)	13,527 (86.2%)	13,507 (86.1%)
QM #0687 Percent of Long-stay Residents Who Were Physically Restrained (Long Stay)	13,864 (88.4%)	13,856 (88.3%)
QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)	12,735 (81.2%)	12,714 (81.1%)
QM #0689 Percent of Residents Who Lost Too Much Weight (Long Stay)	13,761 (87.7%)	13,745 (87.6%)
QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)	13,775 (87.9%)	13,768 (87.8%)

NOTE: Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay residents (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db135_request_v1 \db135_request_v1.log)

3.3 Reliability

Reliability indicates whether changes in QM scores from reporting period to reporting period reflect actual changes in quality as opposed to random fluctuation that results from factors other than quality of care. Dramatic changes in facility QM scores from one quarter to the next may be an indicator of measure instability rather than improvement or decline in quality. In this section, we compare facility QM scores between the third and the fourth quarters of 2011 to examine the patterns in change from one quarter to the next. We evaluated the magnitude of change in QM scores including the percentage of facilities showing large changes in QM score (defined by three standard deviations or more). In an earlier analysis of MDS 2.0 QMs, the University of Colorado (2008) employed an analytic approach that used deciles as its preferred measure of quarter-to-quarter QM change. We applied this approach in MDS 3.0 data. However, considering that small score changes can lead to substantial rank changes for measures with low variability we also examined the QM score change, using each QM score's standard deviation as the normalized unit of change.

Table 3-3 summarizes the results of this analysis. Overall, the measures appear stable: for each QM, at least 71.2 percent (QM #685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder [Long Stay]) and as many as 88.6 percent (QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine [Long Stay]) of facilities had the same score (within one standard deviation) in for the third and the fourth quarters of 2011. Quarter by quarter differences in scores were generally small: the average change in score for all measures was less than 1 percent, except for QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay), on which scores changed by an average of 2.6 percent. Changes of more than three standard deviations were rare: the QM with the greatest proportion of large changes, QM #0681 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Long Stay) had just 2.9 percent of facilities exhibiting large changes in either direction.

Table 3-3
Reliability: proportions of facilities with declining or improving quality measure scores from Quarter 3 to Quarter 4, 2011

Measure name	*n* of facilities reporting	Mean QM score change, %	SD of QM score change, %	Declined by more than three SDs, %	Declined by between two and three SDs, %	Declined by between one and two SDs, %	Declined by less than one SD, %	Improved by less than one SD, %	Improved by between one and two SDs, %	Improved by between two and three SDs, %	Improved by more than three SDs, %
QM #0674 Percent of Residents Experiencing One of More Falls With Major Injury (Long Stay)**	13,729	-0.1	1.9	0.6	2.0	11.0	28.7	44.1	10.8	2.2	0.7
QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)	10,431	-0.2	7.2	0.6	2.2	10.0	37.0	37.0	10.5	2.2	0.5
QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)	11,398	-0.5	5.3	0.7	2.1	10.5	34.9	38.7	10.5	2.2	0.6
QM #0678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay)	10,321	-0.2	2.1	1.0	2.2	7.9	26.7	51.9	7.7	1.9	0.8
QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)	10,939	-0.1	4.0	0.5	2.4	10.8	34.7	37.8	11.2	2.3	0.4
QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay)	9,767	2.6	13.8	0.4	1.6	8.0	50.3	26.7	8.5	3.4	1.2

(continued)

Table 3-3 (continued)
Reliability: proportions of facilities with declining or improving quality measure scores from Quarter 3 to Quarter 4, 2011

Measure name	*n* of facilities reporting	Mean QM score change, %	SD of QM score change, %	Declined by more than three SDs, %	Declined by between two and three SDs, %	Declined by between one and two SDs, %	Declined by less than one SD, %	Improved by less than one SD, %	Improved by between one and two SDs, %	Improved by between two and three SDs, %	Improved by more than three SDs, %
QM #0681 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Long Stay)	135,134	-0.3	18.4	0.3	1.5	7.8	38.5	43.5	4.3	1.5	2.6
QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)	12,405	0.9	8.2	0.8	1.9	7.1	47.1	32.0	7.6	2.4	1.1
QM #0683 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)	13,732	0.1	6.0	0.9	1.0	4.0	61.2	27.4	3.2	1.1	1.2
QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay)	13,620	-0.1	4.5	0.5	2.4	10.3	33.8	39.9	10.2	2.4	0.5
QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)	5,948	0.7	7.2	0.3	2.0	12.3	36.2	35.0	11.1	2.5	0.5
QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (long stay	13,289	-0.2	2.3	0.7	2.4	10.1	32.1	41.8	10.4	1.9	0.7

(continued)

Table 3-3 (continued)
Reliability: proportions of facilities with declining or improving quality measure scores from Quarter 3 to Quarter 4, 2011

Measure name	*n* of facilities reporting	Mean QM score change, %	SD of QM score change, %	Declined by more than three SDs, %	Declined by between two and three SDs, %	Declined by between one and two SDs, %	Declined by less than one SD, %	Improved by less than one SD, %	Improved by between one and two SDs, %	Improved by between two and three SDs, %	Improved by more than three SDs, %
QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)	13,721	-0.1	1.7	1.3	1.9	6.3	14.9	67.0	6.1	1.6	1.0
QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Have Increased (Long Stay)*	12,385	0.2	9.1	0.5	2.0	10.3	38.3	36.5	9.7	2.0	0.8
QM # 0689 Percent of Residents Who Lose Too Much Weight (Long Stay)*	13,601	-0.1	4.2	0.5	2.2	11.0	34.8	38.1	10.7	2.4	0.5
QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)	13,602	0.4	5.3	0.4	1.3	6.1	55.8	28.2	5.4	1.4	1.4

QM = quality measure; SD = standard deviation.

*Measures average QM over 2 quarters of data.

**Measures average QM score over 4 quarters of data.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

3.4 Validity Analysis—Impact of Discharge Assessment on QM Scores

Earlier, we examined the impact of using discharge assessments in the MDS 3.0 on nursing homes' ability to meet minimum sample size requirements for public reporting of the 16 NQF-endorsed QMs. Here, we examine the impact of the inclusion of the MDS 3.0 discharge assessment on QM scores as a validity check. To determine the impact, we removed discharge assessments from the construction of resident episodes, looking back instead to the next eligible assessment for each resident, and recalculating scores for each QM and for each facility based on the resulting set of target assessments. ***Table 3-4*** shows the results of this analysis.

The key finding of this analysis is that discharge assessments tend not to result in large changes in QM scores:

- The largest shift in scores was observed for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay), on which facilities averaged a 3.9 percent decrease in their QM scores when discharge assessments were excluded. However, some caveats apply for this measure.
  - Previous analyses of the MDS 3.0 indicated that discharge assessments, particularly for short-stay residents, tend to show higher levels of missing data on pain items.
  - The results here suggest that either residents self-report less pain during their stay than they do upon discharge or that the high levels of missing data are suppressing the magnitude of scores for this measure. If this level is intolerable, measure developers may consider other strategies to reduce the impact of missing data. This could involve changes in the specification of the measure, such as excluding unexpected discharges (the unplanned discharge was a new assessment item set introduced to the MDS 3.0 by CMS in April 2012) from the specifications of this measure, or pulling values forward from the assessment prior to the discharge. Additional provider training regarding the importance of completing the discharge assessment may be warranted.
- Scores for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) of 1.0 percent after excluding discharge assessments. The remaining 14 measures saw mean changes of less than half of one percent.

Taken together with the results of our reportability analysis, it appears that including discharge assessments in the calculation of QM scores with MDS 3.0 data has a negligible effect on scores for most measures and generally allows more facilities to be included in measure calculation and reporting.

Table 3-4
Facility-level quality measure score changes after excluding discharge assessments

Measure name	*n*	Mean score change, %	SD, %	10th percentile, %	25th percentile, %	50th percentile, %	75th percentile, %	90th percentile, %
QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)**	13,860	-1.0	1.4	-2.9	-1.8	0.0	0	0.0
QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)	10,957	3.9	5.2	-1.7	0.3	3.4	7.0	10.7
QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)+	11,879	0.3	1.5	-1.1	-0.3	-0.0	1.0	2.2
QM #0678 Percent of Residents With Pressure Ulcers That Are New or Have Worsened (Short Stay)+	8,311	-0.1	1.4	-1.7	-0.4	0.0	0.5	1.2
QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)	11,283	-0.1	1.1	-1.3	0.0	0.0	0.1	0.9
QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay)	11,522	-0.3	4.6	-4.8	-2.2	-0.1	1.3	3.9
QM #0681 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Long Stay)	13,779	-0.3	1.7	-2.0	-0.4	0.0	0.0	1.0
QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)	12,578	0.0	3.7	-3.4	-1.2	0.0	1.0	3.3
QM #0683 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)	13,862	0.0	0.9	0.0	0.0	0.0	0.0	0.0

(continued)

Table 3-4 (continued)
Facility-level quality measure score changes after excluding discharge assessments

Measure name	*n*	Mean score change, %	SD, %	10th percentile, %	25th percentile, %	50th percentile, %	75th percentile, %	90th percentile, %
QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay)	13,758	0.0	1.1	-1.3	0.0	0.0	0.1	1.2
QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)	6,541	-0.1	1.8	-2.3	-1.0	0.0	0.7	2.0
QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)+	13,507	0.0	0.7	-0.6	0.0	0.0	0.0	0.3
QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)	13,856	0.0	0.3	0.0	0.0	0.0	0.0	0.0
QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Have Increased (Long Stay)*	12,714	0.0	1.7	-2.1	-0.8	0.0	0.6	2.0
QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)*	13,745	0.0	1.0	-1.1	0.0	0.0	0.1	1.0
QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)	13,768	0.0	0.9	-0.7	0.0	0.0	0.0	0.9

NOTE: *n* = number of facilities able to report the QM both before and after discharge assessments were excluded.

SD = standard deviation; QM = quality measure.

*Measures average QM over 2 quarters of data.

**Measures average QM score over 4 quarters of data.

+Model-adjusted risk adjustment applied.

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request_v1\db135_request_v1.log)

3.5 Validity Analysis—Variation by State

For a QM to be valid, variation observed in the distribution of the QM should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as state-level payment policies or demographics, this can be a threat to the validity of the measure. To explore whether state characteristics might be a source of facility score variation we conducted a one-way analysis of variance (ANOVA) and reported the proportion of variance in QM scores accounted for by state. We also examined the interquartile difference between the mean state-level scores for states.

Table 3-5 summarizes the results of the analyses of state variation for each QM. The proportion of variance in each QM explained by the state that facilities are located varies across QMs, ranging from 2.0 percent for QM #0678: Percent of residents with pressure ulcers that are new or worsened (Short Stay) to 16.2 percent for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay). It is possible that state explains a larger proportion of variance in QMs constructed by items also used in Resource Utilization Groups Version 4 (RUG IV—used to calculate Medicare and/or Medicaid payment to nursing homes) as there is state variation in implementation of the RUG system for Medicaid reimbursement adjustment (http://ltcfocus.org/). For example, MDS 3.0 depressive symptoms items used to calculate the QM are also used as supporting items in RUG (e.g., triggers for skip patterns, none-of-the-above items and component item for summary score). For QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay), facility state location accounts for 11.2 percent of the variance in the QM score. However, the pattern is not substantiated in the MDS ADL items used in RUG IV; facility state location explains only 6 percent of the variance in QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay). Overall, the facility state location explains a small proportion of variance in most of QMs. For QMs that vary to a greater extent by state, further examination of the cause of the variation may be necessary prior to proposing a policy remedy.

Table 3-5
Proportion of variance attributed to state-by-state differences for quality measures

Measure name	*n* of states/territories reporting	*F* ratio	Proportion of variance accounted for by state (η^2)	Interquartile range
QM #0674 Percent of Residents Experiencing One of More Falls With Major Injury (Long Stay)	52	22.9*	7.8%	1.1%
QM #0676: Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)	53	16.5*	7.2%	6.4%
QM#0677, Percentage of Residents Who Self-Report Moderate to Severe Pain (Long Stay)	51	24.9*	9.5%	3.5%
QM #0678: Percent of residents with pressure ulcers that are new or worsened (Short Stay)	53	4.3*	2.0%	0.5%
QM #0679, Percent of High-Risk Residents With Pressure Ulcers (Long Stay)	51	15.6*	6.5%	2.1%
QM #0680: Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)	53	9.5*	4.0%	4.6%
QM #0681 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (long-stay)	52	9.2*	3.3%	4.2%
QM #0682: Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)	53	12.3*	4.8%	6.8%
QM #0683, Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)	52	12.7*	4.5%	3.6%
QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay)	52	10.8*	3.9%	1.8%
QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)	51	25.6*	16.2%	7.6%

(continued)

Table 3-5 (continued)
Proportion of variance attributed to state-by-state differences for quality measures

Measure name	*n* of states/territories reporting	*F* ratio	Proportion of variance accounted for by state (η^2)	Interquartile range
QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)	51	11.2*	4.0%	1.5%
QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)	52	20.0*	6.9%	1.4%
QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)	51	18.7*	6.0%	3.3%
QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)	51	9.3	3.3%	1.3%
QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)	52	34.0	11.2%	3.9%

QM = quality measure.

NOTE: *$p < 0.001$

Analysis Date: 6/28/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

3.6 Summary of Risk-Adjustment Strategies

Risk adjustment seeks to minimize the influence of factors beyond the control of nursing facilities on QM scores and to maximize the likelihood that these measures accurately assess facility quality. Risk adjustment may also be employed to ensure that measures compare care for patients at similar risk across facilities. For the 16 NQF-endorsed QMs discussed in this report, there are two primary methods of risk adjustment: model-based adjustment and denominator exclusion or sample restriction.

Denominator exclusion and sample restriction share a similar logic that they both can make the denominator a more homogenous group of residents with regard to their risk for the QM outcome. An example of denominator exclusion is QM #0688 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay), which excludes residents who are unable to self report pain because pain ascertainment has been shown to be less effective when using staff assessment than self-report (Saliba & Buchanan, 2008). Sample restriction uses MDS 3.0 items to identify residents as either high- or low-risk based on conditions relevant to a given QM. For example, QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) only includes residents at high-risk for acquiring pressure ulcers based on three criteria—impaired bed mobility or transfer, comatose or malnutrition or at risk of malnutrition. This approach can both minimize floor effects (i.e., many facilities with a low prevalence) and, more importantly for risk-adjustment purposes, allow for an examination of comparable residents across different facilities.

Model-based risk adjustment works on a heterogeneous group of residents but takes the heterogeneity into account when calculating the QM scores. For this approach, MDS items indicating certain health conditions that increase or decrease the likelihood of a health outcome are identified as covariates. A logistic regression model is fitted and the estimated coefficients are used to predict the probability that a resident will experience that outcome given the covariate values. The average of these resident-level probabilities for a given facility represents that facility's expected score for that measure. The final QM score for a facility is found by combining its observed score (i.e., the prevalence or incidence of the outcome) with its expected score.

Tables 3-6 and ***3-7*** provide a summary of risk-adjustment techniques used with four NQF-endorsed short-stay QMs and 12 NQF-endorsed long-stay QMs, respectively. The tables delineate which QMs are associated with risk-adjustment techniques, which specific techniques are used for those measures, and which covariates are used in model-based risk adjustment (where applicable). ***Tables 3-8*** and ***3-9*** identify sample exclusions for short-stay and long-stay measures, respectively, as well as the rationale for selecting certain items as exclusion criteria.

Table 3-6
Summary of MDS 3.0 quality measures and risk-adjustment techniques used (short stay)

Quality measure number	Quality measure	Exclusion	Model-based	Covariates used
0676	Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)	Yes	No	Not applicable.
0678	Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)	No*	Yes	▪ Limited or more assistance in bed mobility self-performance ▪ Bowel incontinence at least occasionally ▪ Diabetes or Peripheral Vascular Disease ▪ Low BMI (<19)
0680	Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)	Yes	No	Not applicable.
0682	Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay)	No	No	Not applicable.

*Exclusions are for missing or illogical data

Table 3-7
Summary of MDS 3.0 quality measures and risk-adjustment techniques used (long stay)

Quality measure number	Quality measure	Exclusion	Model-based	Covariates used
0674	Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)	Missing data only	No	Not applicable.
0677	Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)	Yes	Yes	Independence/ modified independence in daily decision making
0679	Percent of High-Risk Residents With Pressure Ulcers (Long Stay)	Yes	No	Not applicable.
0681	Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)	Yes	No	Not applicable.
0683	Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)	No	No	Not applicable.
0684	Percent of Residents with a Urinary Tract Infection (Long Stay)	No*	No	Not applicable.
0685	Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)	Yes	No	Not applicable.

(continued)

Table 3-7 (continued)
Summary of MDS 3.0 quality measures and risk-adjustment techniques used (long stay)

Quality measure number	Quality measure	Exclusion	Model-based	Covariates used
0686	Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)	Yes	Yes	Frequent bowel incontinence Pressure ulcer stage 2, 3, or 4
0687	Percent of Residents Who Were Physically Restrained (Long Stay)	Missing data only	No	Not applicable.
0688	Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)	Yes	No	Not applicable.
0689	Percent of Residents Who Lose Too Much Weight (Long Stay)	Missing data only	No	Not applicable.
0690	Percent of Residents Who Have Depressive Symptoms (Long Stay)	Yes	No	Not applicable.

*Exclusions are for missing data or assessment type

Table 3-8
Summary table of quality measures with denominator exclusion/sample restriction used and rationale (short stay)

Quality measure number	Quality measure	Denominator exclusion/sample restriction
0676	Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)	Resident must be able to report pain
0678	Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)	Exclusions are due to missing or illogical data, not risk adjustment
0680	Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)	Resident excluded if not in facility during the current or most recent influenza season
0682	Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay)	Not applicable

Table 3-9
Summary table of quality measures with denominator exclusion/sample restriction used and rationale (long stay)

Quality measure number	Quality measure	Denominator exclusion/sample restriction
0674	Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)	Exclusions are associated with missing data, not risk adjustment
0677	Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)	Resident must be able to report pain
0679	Percent of High-Risk Residents With Pressure Ulcers (Long Stay)	Must meet definition for high risk which is defined as one or more of the following: impaired bed mobility or transfer; comatose; or malnutrition or at risk for malnutrition
0681	Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)	Resident was not in facility during the current or most recent influenza season
0683	Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)	Not applicable
0684	Percent of Residents with a Urinary Tract Infection (Long Stay)	Exclusions are for data availability or assessment type, not risk adjustment
0685	Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)	Residents who have any of the following high risk conditions ▪ Severe cognitive impairment ▪ Totally dependent in bed mobility self-performance ▪ Totally dependent in transfer self-performance ▪ Totally dependent in locomotion on unit self-performance ▪ Resident is comatose ▪ Resident has indwelling catheter (or status is missing) ▪ Resident has ostomy (or status is missing)
0686	Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)	Target assessment indicates neurogenic bladder (or status is missing) or obstructive uropathy (or status is missing)
0687	Percent of Residents Who Were Physically Restrained (Long Stay)	Exclusions are for data availability or assessment type, not risk adjustment
0689	Percent of Residents Who Lose Too Much Weight (Long Stay)	Exclusions are for data availability or assessment type, not risk adjustment
0690	Percent of Residents Who Have Depressive Symptoms (Long Stay)	Resident is comatose (or status missing)

3.7 Summary

In Section 3, we presented results for all 16 NQF-endorsed measures on several aspects important for evaluating QMs: variability, reportability, reliability, and validity (operationalized as missing data rates; we examined correlations of paired and related measures in the QM specific sections are listed in Section 4). We also examined the impact of including the new MDS 3.0 discharge assessment in the specifications of QMs.

To assess variability, we examined distributions of facility-level QM scores, focusing on the magnitude of interquartile ranges and proportions of perfect scores. If scores for a particular measure cluster together in a narrow range, then that measure may not be able to detect meaningful differences in facility performance. The same is true for measures with clusters of scores at the top of the range or perfect scores. Variability of QM scores, as measured by interquartile range, is highest for QM #0585 Percent of Long-stay, Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay); QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay); and QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay). Measures with low prevalence or incidence also tended to have narrow interquartile ranges and high proportions of perfect scores; for example, QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) and QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay). Large proportions of perfect scores were also seen for the long-stay vaccination measures. Thus, measures with prevalence or incidence rates at the extreme ends of the range of possible scores merit careful monitoring to see if they can capture meaningful differences in facility quality given the relatively small variability. These measures, however, can be useful in identifying facilities that perform very poorly in these aspects of care.

Analyses examining the impact of the new MDS 3.0 discharge assessment on reportability and on the distribution of scores for the 16 QMs generally indicate that the new item set has only modest effects. CMS requires that nursing homes have 20 eligible residents to report on short-stay measures and 30 eligible residents to report on long-stay measures. Removing the discharge assessment from the measure specifications for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) led to a significant reduction in the number of facilities with sufficient sample size, thus, adding discharge assessments provides a substantial boost in reportability for that measure. However, for the other measures, the effect was far more modest (a difference of fewer than 3 percent of facilities for each measure). Similarly, discharge assessments appeared to have little effect on measure scores. The largest shift was seen for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay), on which scores decreased by an average of 3.9 percentage points. While this is a relatively small change, this difference does call attention to high rate of missing data on pain-related items on discharge assessment; an issue that warrants further investigation. All other measures saw even smaller changes (none greater than 1.0 percent). Thus, the new discharge assessment appears to have some positive effects on reportability while not meaningfully altering mean QM scores.

Overall, the 16 measures discussed here appear to be reliable over time. We examined the two most recent quarters of MDS 3.0 data. For 15 of the 16 measures, the average change in

QM score across facilities between quarters was less than 1 percent (the 16th saw a change of 2.6 percent). For each QM, the vast majority of facilities received approximately the same score (within one standard deviation). Large differences (defined by three standard deviations) in scores from one quarter were exceedingly rare. QM #0681 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Long Stay) showed the greatest proportion of facilities whose score changed by more than 3 standard deviations (2.9 percent). This is also one of the measures with the greatest variability.

Thus, the 16 NQF-endorsed QMs discussed in this report perform well in terms of reportability and reliability. The addition of discharge assessments to measure specifications had little effect on national mean QM scores but provided a large increase in reportability for one measure and small increases in reportability for most. Some QMs with score values close to perfect scores tend to have smaller variability (e.g., the mean score for QM #0683 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination [Long Stay] is 93.8 percent and the interquartile range is 6.4 percent). However, most of the QMs show good variability. Missing data does not appear to have a substantial impact across measures. We reported the proportion of variance in QM score explained by state to examine the extent to which the variance is attributable to factors outside nursing homes' control. For most QMs, the state that facilities are located explains only a small proportion of the variance in score. These analyses will be discussed in greater and detail in Section 4 along with other, measure-specific analyses.

SECTION 4
QUALITY MEASURE–SPECIFIC ANALYSES

4.1 QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

4.1.1 Summary of Findings

Quality Measure (QM) #0676 Percent of Residents who Self-Report Moderate to Severe Pain (Short Stay) reports the percentage of short-stay residents who self-report daily pain with at least one episode of moderate/severe pain or very severe/horrible pain of any frequency. This section presents descriptive analyses and test results for this QM. Of all short-stay resident episodes, 71.5 percent meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 70.3 percent of the facilities are able to report this QM. Among facilities able to report, the mean facility QM score is 23.1 percent. The QM score varies across facilities with a standard deviation of 12.6 percent and an interquartile range of 17.2 percent, suggesting acceptable variability in the measure's ability to differentiate among facilities with poor and good quality of care.

The introduction of the discharge assessment and its inclusion as a target assessment for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) is a new feature of Minimum Data Set (MDS) 3.0 and MDS 3.0 QMs. Our analyses indicated that the items used to calculate QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) often have missing values on the discharge assessment. Assessments with missing item values, and thus resident episodes, are dropped from the calculation of the measure, affecting QM reportability. Analyses illustrated that including discharge assessments in the set of target assessments, decreases the resident episodes in the denominator by about 8 percent and the percent of facilities able to report this QM by approximately 5 percent.

As reliability checks of this QM, we tracked changes in QM score and in percentile ranking by quarter. The majority of facility score changes varied only one standard deviation, with a very low proportion of facility QM score changes greater than three standard deviations. A similar pattern was found for facility rank changes. About two-thirds of facilities remained within the same decile from quarter to quarter. Almost all facilities had rank changes within three deciles. These findings demonstrate acceptable reliability for this QM.

Several validity tests were conducted for this QM. First, the correlations between QMs within a "care process group" were analyzed. The hypothesis was that changes in one QM within a measure group (e.g., short-stay and long-stay pain QMs) should be correlated with changes in other measures because they reflect similar care processes. Findings illustrated statistically significant and moderate correlations between the short-stay and long-stay pain QMs ($r = 0.560$; $p < 0.001$). Second, we examined the frequency and distribution of missing data. We also found that missing rate only varies slightly by resident characteristics (e.g., age, gender, and cognitive status). The missing rate also varies across facilities and was higher among facilities with higher values in the QM score (indicating poorer quality). Finally, we analyzed the potential geographic (state) and seasonal (quarterly) variations in this QM. The state in which

the facility is located explains about 7 percent of the variation in this QM. The QM score remains largely stable from Quarter 1 to Quarter 4 in 2011.

Reliability and validity tests indicate acceptable rigor, and QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) received full National Quality Forum (NQF) endorsement on August 1, 2012.

4.1.2 Background and Introduction to QM

This measure reports the percentage of short-stay residents who have reported almost constant or frequent pain and at least one episode of moderate to severe pain, or any severe or horrible pain, in the 5 days prior to the target assessment. Target assessments may be discharge; 5-, 14-, 30-, 60-, 90-day, or readmission/return Prospective Payment System (PPS) assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

A. Summary of Background for This QM

Unrelieved pain is a source of adverse psychological and physiological effects including nausea, dyspnea, impaired immune response, anxiety, and depression among others (Hanson, 1997; Sachs, Shega, & Cox-Hayley, 2004; Scherder & Bouma, 2000; Wrede-Seaman, 2001). Failure to identify pain can result in lower quality of life for nursing home residents—both because it can be a symptom of an underlying medical problem, and because untreated pain can result in decline of functional ability. This is especially the case in residents with cognitive impairment. Unrelieved pain costs millions of dollars annually as a result of longer hospital stays, increased rehospitalizations, increased utilization of outpatient care, and emergency room visits (Grant, Ferrell, Rivera, & Lee, 1995; Wu, Miller, Lapane, & Gozalo, 2003). Pain management is of great interest to the public as evidenced by being one of the major goals of the Advancing Excellence in America's Nursing Homes initiative (http://www.nhqualitycampaign.org/). Pain management in nursing homes is central to the Omnibus Budget Reconciliation Act of 1987 (OBRA '87) mandate to promote the "maximum practicable functioning" of residents. Failure to identify and address pain denies residents the right to freedom from neglect (Wiener, Freiman, & Osterweil, 2007). Evidence suggests that at least 40 to 85 percent of nursing home residents have persistent pain, and that pain is often not fully documented and therefore the percentage may be even higher (Centers for Medicare & Medicaid Services, 2008; Ferrell, Ferrell, & Osterweil, 1990; Mor, Zinn, Angelelli, Teno, & Miller, 2004; Parmelee, Smith, & Katz, 1993; Sengstaken & King, 1993; Weiner & Rudy, 2002; Wu et al., 2003). A standard measure to reflect the quality of care related to pain assessment and treatment provides a benchmark for pain management practices that vary widely across nursing homes.

The MDS 3.0 pain QMs are now based on patient interview or staff assessment if the patient is determined not to be able to self-report pain. The change to interview-based items was made because research demonstrated that self-reported severity and frequency of pain using standardized scales is significantly more accurate than staff assessment of pain. However, one concern with the self-reporting of pain is that it may be underrepresented in patients with cognitive impairment. One study of nursing home resident use of analgesics for pain relief found that the receipt of both opioid and nonopioid analgesics was significantly associated with the

ability to self-report pain, and that residents who received analgesics had higher scores on the Mini Mental Status Exam (MMSE) indicating better cognitive functioning than those who were prescribed as-needed analgesics that were never administered (Allen et al., 2003). Analysis of the patients' medical records showed that although 86 percent of patients had been diagnosed with a painful condition, only 70 percent received analgesics. Based on the results of the study, the authors concluded that residents with impaired cognitive status were less able to self-report pain and less likely to receive analgesics for pain relief (Allen et al., 2003). There are methods for identifying pain in patients with cognitive impairment including the use of scales and assessments to determine the presence of symptoms such as behavioral problems or depression (Leone, Standoli, & Hirth, 2009), or behaviors such as crying, yelling, tactility, wincing, or restricted movement (Closs, Cash, Barr, & Briggs, 2005). The MDS 3.0 items for identifying cognitive impairment are based on the Brief Interview for Mental Status (BIMS). BIMS was tested and showed a sensitivity of 69.7 to 94.4 percent and a specificity of 85.6 to 78.6 percent for severe cognitive impairment, indicating that the majority of patients with cognitive impairments are being correctly identified (Chodosh et al., 2008). BIMS showed high reliability as well, with facility nurses and research assistants obtaining identical mean scores (Chodosh et al., 2008).

B. Summary of Differences between MDS 2.0 and MDS 3.0 QM Definitions

QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) reports the percentage of short-stay residents with a target assessment during the selected quarter who report almost constant or frequent pain and at least one episode of moderate to severe pain, or any severe or horrible pain, in the 5 days prior to the target assessment. The numerator is the number of short-stay residents who are able to self-report (item J200=1), who have a 14-day PPS assessment during the preceding 6 months, who report almost constant or frequent pain (item J0400 = 1 or 2) and at least one episode of moderate to severe pain (item J0600A = 5, 6, 7, 8, or 9 on a scale of 1–10, with 10 being the worst pain you can imagine, or item J0600B = 2 or 3 on a scale of 0–4, with 4 being very severe, horrible pain) or very severe/horrible pain of any frequency (item J0600A = 10 on a scale of 1 to 10 or item J0600B = 4 on a scale of 0 to 4) in the 5 days prior to the 14-day PPS assessment. The denominator includes all short-stay residents in the nursing facility with a target assessment during the preceding 6 months from the selected quarter who do not meet the exclusion criteria. Residents are excluded from the denominator if they did not meet the pain symptom conditions for the numerator and any of the following conditions is true: the numerator pain intensity item indicates no pain (J0600A=0); the resident cannot self-report; or there is missing data in the responses to the relevant questions on the MDS assessment. If the facility QM calculation includes fewer than 20 residents in the denominator (after exclusions are applied), then the facility is excluded from public reporting due to small sample size.

This measure was updated from the MDS 2.0 QM in several ways to address shortcomings identified by the DAVE 2 project (assessed reliability between MDS 2.0 and MDS 3.0 items/QMs; Abt Associates Inc., Stepwise Systems Inc., & Qualidigm, 2007) and the empirical review conducted by the University of Colorado as part of the DMINHo Project (evaluated MDS 3.0 QM; Brega, Goodrich, Nuccio, et al., 2008; Brega et al., 2007). Specifically, the DMINHo Project found the measure was reportable for 75.6 percent of facilities providing post-acute care, and demonstrated acceptable variability among facilities. However,

the measure was found to perform poorly on indicators of validity, stability, reliability, and risk adjustment. According to their report, 11.1 percent of facilities experienced changes in rankings of three deciles or more, and for the majority of the time, these changes didn't reflect actual changes in quality but rather erratic changes in facility performance (i.e., changes in QM score that did not demonstrate consistent improvement or decline over time). In addition, research showed that the MDS 2.0 measure underestimated the prevalence of pain (Saliba & Buchanan, 2008). The MDS 3.0 QM changed the look-back period from 7 days to 5 days. In addition, previously the MDS 2.0 measure was based entirely on the 14-day PPS assessment. Finally, the QM is now based on resident interview items, using numeric or verbal descriptor scales to assess frequency and intensity of pain. The change to interview-based items was made because research has demonstrated that self-reported severity and frequency of pain using standardized scales is significantly more accurate than staff assessment of pain. An interview-based measure also aligns pain assessment with accepted care standards across settings.

Note also that the overall definitions changed from MDS 2.0 to MDS 3.0. For MDS 2.0, residents were included in the post-acute care measures if they had a 14-day PPS MDS in the target quarters, and measures were calculated based on just 14-day PPS assessments. For MDS 3.0, the analogous short-stay resident population is defined as residents with 100 or fewer cumulative days in facility, and assessments may be discharge, 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments, or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections summarize the basic analyses performed on QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) using MDS 3.0 episode file for the QMs for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- Number of assessments that are included and excluded from the numerator of the QM based on sample restrictions
- Whole sample prevalence of items used to construct QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)
- Findings regarding QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) variability, reportability, reliability, and validity

4.1.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, 987,692 (71.5 percent) assessments were included in the denominator of QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay), and a total of 393,336 resident episodes (28.5 percent of the total number of short-stay resident episodes) were excluded from the denominator. ***Table 4.1-1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criterion. Note that one resident episode can meet multiple exclusion criteria. The number and percent of resident episodes that met each exclusion criterion are as following:

- A total of 176,137 (12.8 percent) resident episodes met exclusion criterion 1 (missing data)
- 216,641 (15.7 percent) resident episodes met exclusion criterion 2 (unable to participate in pain interview);
- 312 (0.0 percent) assessments met exclusion criterion 3 (inconsistent: J0300 (any pain or hurting at any time in the last 5 days) = [1] and J0600A (pain intensity—numeric rating scale) = [00]).

Table 4.1-1
Target assessments used in the calculation of QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	987,692	71.5%
Excluded—Exclusion 1	176,137	12.8%
Excluded—Exclusion 2	216,641	15.7%
Excluded—Exclusion 3	312	0.0%
Total number of short-stay resident episodes	1,381,028	—

NOTES:

Exclusion 1 = Missing data

Exclusion 2 = Unable to participate in pain interview

Exclusion 3 = Inconsistent: J0300 (any pain or hurting at any time in the last 5 days) = [1] and J0600A (pain intensity—numeric rating scale) = [00]

[a] Column values may not add up to total since a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident assessments in short-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

Table 4.1-2 shows the frequency and percentage responses on items used to define the numerator for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay). Item J0400 asks the resident about the frequency of pain in the last 5 days. The response categories are 1 (almost constantly) to 4 (rarely). Item J0600A asks the resident about pain intensity using a 0 (no pain) to 10 (worst imaginable pain) scale. Item J0600B asks the resident about pain using a verbal descriptor scale with a range of 1 (mild pain) to 4 (very severe, horrible). For all three items, a code of 9 indicates the resident was unable to answer. Overall, the triggering rates were distributed across all scores, and did not demonstrate a floor or ceiling effect. The table shows that the frequency of responses to item J0400 is greatest for responses of 2 and 3, indicating that the majority of residents triggering the QM report frequent (2) or occasional (3) pain in the past 5 days. On Item J0600A, the highest frequencies were observed for resident responses of 5 and 6 (corresponding to moderate pain) on a 10-point pain scale.

Similarly, on J0600B, the highest frequency was seen for resident response 2 (moderate pain). This same pattern was also noted for the long-stay pain measure, QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay), though the percentages of residents reporting pain at any level were notably higher for the short-stay population than for the long-stay population. This may be in part because short-stay residents are more likely to be receiving rehabilitation services such as physical therapy for conditions such as hip or knee replacements where break through pain may occur, even with regularly scheduled pain regimen in place. Also notable is the skip frequency for each pain item used to calculate the measure. A caret (^) indicates a skip; for QM #0676 items, a skip is based on response to Item J0300 ("have you had pain or hurting at any time in the last 5 days?"). Items J0400, J0600A, and J0600B are skipped for roughly 47 percent, 56 percent, and 73 percent of residents respectively. These numbers are lower than for the long-stay pain measure.

Table 4.1-2
Frequency and percentage of responses on items used to define the numerator for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

Item	Entry	Frequency	Percentage	Cumulative %
J0400	— = missing	164,539	11.9%	11.9%
J0400	1 = Almost Constantly	55,841	4.0	16.0
J0400	2 = Frequently	198,529	14.4	30.3
J0400	3 = Occasionally	256,576	18.6	48.9
J0400	4 = Rarely	48,343	3.5	52.4
J0400	9 = Unable to Answer	4,067	0.3	52.7
J0400	^ = skipped	653,133	47.3	100.0
J0400	Total	1,381,028	100.0	—
J0600A	— = missing	171,100	12.4	12.4
J0600A	0	663	0.1	12.4
J0600A	1	3,945	0.3	12.7
J0600A	2	20,895	1.5	14.2
J0600A	3	40,591	2.9	17.2
J0600A	4	55,075	4.0	21.2
J0600A	5	77,711	5.6	26.8
J0600A	6	69,050	5.0	31.8
J0600A	7	54,413	3.9	35.7
J0600A	8	67,270	4.9	40.6
J0600A	9	17,861	1.3	41.9
J0600A	10	26,641	1.9	43.8
J0600A	99	7,101	0.5	44.4
J0600A	^ = skipped	768,712	55.7	100.0
J0600A	Total	1,381,028	—	—
J0600B	— = missing	183,922	13.3%	13.3%
J0600B	1 = Mild	51,334	3.7	17.0
J0600B	2 = Moderate	107,602	7.8	24.8

(continued)

Table 4.1-2 (continued)
Frequency and percentage responses on items used to define the numerator for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

Item	Entry	Frequency	Percentage	Cumulative %
J0600B	3 = Severe	29,289	2.1	27.0
J0600B	4 = Very severe, horrible	4,212	0.3	27.3
J0600B	9 = Unable to answer	3,178	0.2	27.5
J0600B	^ = skipped	1,001,491	72.5	100.0
J0600B	Total	1,381,028	—	—

NOTE: J0400 = Pain Frequency; J0600a = Numeric Rating Scale; J0600b = Verbal Descriptor Scale

Analysis Date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_010_10.log)

4.1.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores. This can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles) or shown by the percentage of nursing homes with "perfect scores" (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes) which may indicate a "ceiling effect". ***Table 4.1-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0676 is 23.1 percent with a standard deviation of 12.6 percent. This QM has an interquartile range of 17.2 percent. About 1.4 percent of facilities have scores of 0 percent ("perfect" scores).

Table 4.1-3
QM score distribution for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
11,033	23.1%	12.6%	7.4%	13.8%	22.2%	31.0%	39.5%	1.4%	17.2%

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure

QM scores are reported at the facility-level

Analysis Date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_010_10.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 20 residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examine the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.1-4*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 70.3 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) using MDS 3.0. This is an improvement over the MDS 2.0 measure, for which 68.9 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure, increasing the number of resident episodes eligible to be included in the QM calculation. The inclusion of the discharge assessment captures short-stay residents who are discharged prior to their 14-day assessment who would have been missed in the MDS 2.0 specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, slightly more than 75 percent of target assessments for the short-stay population for Quarter 4, 2011, were discharge assessments compared to 8.8 percent for the long-stay population. This suggests that the inclusion of the discharge assessment may impact the reportability of short-stay QMs to a greater degree than the long-stay QMs.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay), we report changes in the number of facilities able to report this QM when discharge assessments are not included among target assessments. The results are presented in ***Table 4.1-4***. Excluding discharge assessments among target assessments results in many more residents being included in the calculation of the measure and a corresponding increase in QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) reportability: increasing from 70.3 percent (with discharge assessments) to 75.6 percent (without discharge assessments). The increase in reportability after excluding discharge assessments from the set of eligible target assessments may suggest that discharge assessments have a higher rate of missing value on pain-related items than other PPS assessments do (see also ***Table 4.1-7***). Missing data on discharge assessments may indicate an unexpected discharge or be evidence that nursing homes have not yet incorporated completing clinical information on the discharge assessment into their routine care process. Regardless of the reason, additional training of providers regarding the discharge assessment may be warranted.

Table 4.1-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (Percent) of resident episodes in the denominator	987,692 (71.5%)[a]	1,094,421 (79.2%)[a]
Number (Percent) of resident episodes in the numerator	232,454 (16.8%)[a]	303,434 (22.0%)[a]
Number (Percent) of facilities able to report this QM	11,033 (70.3%)[b]	11,851 (75.6%)[b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the short-stay population (1,381,028).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay residents (15,686).

Analysis date: 5/29/2012

SOURCE: MDS 3.0, RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\nh_010_10_nd and nh_010_10.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.1-5 illustrates quarterly change in facility rank in 2011. For each pair of quarters, large changes in facility percentile rankings were rare: at least two-thirds of facilities stayed at approximately the same rank, and relatively few facilities (< 6 percent) illustrated rank changes over three deciles.

Table 4.1-5
Distribution of change in facility rank from one quarter to the next, QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	6,866 (66.6%)	6,965 (67.1%)	7,172 (68.8%)
Between 1 and 2 deciles	1,836 (17.8%)	1,890 (18.2%)	1,823 (17.5%)
Between 2 and 3 deciles	917 (8.9%)	908 (8.8%)	835 (8.0%)
More than 3 deciles	695 (6.7%)	611 (5.9%)	601 (5.8%)
Total	10,334 (100.0%)	10,374 (100.0%)	10,431 (100.0%)

Analysis date: 6/15/2012

NOTES:

Total ns reflect facilities that meet minimum requirements for public reporting this QM in both Quarters.

SOURCE: MDS 3.0, RTI analysis of MDS 3.0 episode files for Quarter 1–Quarter 4, 2011(\quarter_4_5\db155_request\db155_request.log)

Table 4.1-6 presents change in facility scores between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is inappropriate to compare absolute changes in facility scores across measures. However, the magnitude of change in an individual measure can be assessed relative to its standard deviation (standard deviation = 12.6 percent). Toward this end, we examined the proportion of facilities whose scores remained approximately the same (score changed less than one standard deviation) and those that changed by more than one, two, or three standard deviations. Overall, the average quarterly change in facility score was a 0.2 percentage point decrease. The vast majority of facilities (74 percent) illustrated less than one standard deviation in their score between the two quarters, while slightly more than 1 percent of the facilities illustrated changes of more than three standard deviations.

The items that make up this measure and the values counted as missing that result in a resident being excluded from the measure are as follows: "Should Pain Assessment Interview be conducted (J0200, [-]), pain presence in the last 5 days (J0300, [9,-]), pain frequency (J0400, [9,-]), and for pain intensity, both the numeric and rating scale items must be missing or invalid responses (J0600A, [99,-] and J0600B, [9,-]). Combined, missing responses across all items comprising this measure resulted in a total of 176, 137 (12.8 percent) of short-stay residents excluded.

Table 4.1-6
Change in facility score from one quarter to the next, QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
10,431	-0.2%	7.2%	0.5%	2.2%	10.0%	37.0%	37.0%	10.5%	2.2%	0.5%

NOTES:

Number of facilities reporting (10,431) reflects facilities that meet minimum requirements for public reporting this QM in both Quarter 3 and Quarter 4, 2011.

Analysis Date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated). For example, the "pain group" of measures, QM # 0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) and QM # 0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay), should reflect the quality of related clinical care processes. Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on related measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically-related group. The analyses are based on facilities' risk-adjusted measures where applicable. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0676 and the facility's percentile rank on QM #0677 for Quarter 4, 2011. We found that the correlation between these two QMs was moderate ($r = 0.560$) and statistically significant ($p < 0.001$).[6] The moderate but significant correlation, as opposed to a strong correlation, may be due to the differences in population (evidence that reporting pain quality measures separately for short- and long-stay residents is appropriate). Care processes for short-stay patients may also be different enough from that of long-stay patients that correlations between the two QMs remain moderate.

Variation by State. For a quality measure to be valid, variation observed in the distribution of the QM should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as state-level payment policies, this can be a threat to the validity of the measure. To explore the question of whether state characteristics might be a source of facility score variation for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) we conducted a one-way analysis of variance (ANOVA). We also examined the interquartile difference between the mean state-level scores across states. The proportion of variance in this measure explained by the state in which facilities are located is 7.2 percent and significant [$F(52, 10{,}980) = 16.5$, $p < .001$].[7] The difference between the mean state-level scores for states at the 25th percentile and the 75th percentile is 6.4 percentage points. Thus, while the majority of the variance in QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) scores is due to factors other than geography, there is a small and significant proportion of the variance that is explained by the respective states in which nursing facilities are located. The reason for the state variation warrants future examination.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns emerge associated with resident types and assessment types, the integrity of

---

[6] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

[7] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

the measure is compromised, that is, the QM is biased and potentially inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

In the fourth quarter of 2011, a total of 393,336 residents were excluded from this QM based on the measure denominator exclusions (i.e., 28.5 percent of the total number of short-stay resident episode). Of these, 216, 641 (15.7 percent) residents were excluded because they were unable to participate in the pain assessment interview (i.e., they were rarely or never understood (J0200 = [0]) or they were unable to answer when asked, "Have you had pain or hurting at any time in the last 5 days? (J0300 = [9]). An additional 312 residents (less than one tenth of 1 percent) were excluded because data on pain items were inconsistent (item J0300 indicated pain or hurting but item J0600A indicated a pain rating of [0]).

For this measure, we analyzed whether missing data on pain items J0300, J0400, or J0600 varied systematically on the following characteristics: age greater than or equal to 85, gender, a score on BIMS less than or equal to 12 (scores indicating cognitive impairment), and inability to make oneself understood (from item B0700). ***Table 4.1-7*** summarizes the results of this analysis. For this measure, age, gender, and cognitive impairment appear to have only minimal effects on rates of missing data for pain items. The inability to make oneself understood appears to have some effect: residents with missing data on these items are more than 4 times as likely to have problems with communication (1.7 percent vs. 0.4 percent). As this measure relies on the ability of a resident to report their own pain, an inability to make oneself understood might naturally lead to difficulty for a facility in properly completing these items on the target assessment. However, this is a relatively low-prevalence condition, and the effect may be magnified by the small size of the percentages in play.

Table 4.1-7
Missing data compared across selected resident characteristics

Resident characteristics	Any missing data on J0300, J0400, or J0600 *n*	Any missing data on J0300, J0400, or J0600 %[a]	No missing data on J0300, J0400, or J0600 *n*	No missing data on J0300, J0400, or J0600 %[b]
Age ≥ 85	51,043	29.0%	303,105	30.7%
Male	70,108	39.8	368,150	37.3
BIMS <= 12	19,506	30.4	265,463	30.3
Inability to make oneself understood (B0700 = 3)	2,754	1.7	3,390	0.4

NOTES:

[a] Percentage reflects proportion of all target assessments with any missing data on J0300, J0400, or J0600 and no missing data on BIMS or B0700.

[b] Percentage reflects proportion of all target assessments with no missing data on J0300, J0400, or J0600 and no missing data on BIMS or B0700.

J0300 = Pain Presence; J0400 = Pain Frequency; J0600 = Pain Intensity

Analysis date: 6/28/2012

SOURCE: RTI Analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db157_request\db157_request.log)

Table 4.1-8 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. The distribution of facility-level missing data rates on items used to construct this measure are positively skewed, with a mean of 12.0 percent and a median of 5.0 percent.

Table 4.1-8
Distribution of facility-level missing rate for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,631	12.0%	16.5%	0%	0%	5.0%	16.7%	35.8%

NOTES:

n = number of facilities who have data for the numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 6/28/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db158_request\db158_request.log)

We also examined the relationship between missing data and QM scores. ***Table 4.1-9*** shows the distributions of facility-level missing rates stratified by quartile of QM scores for this measure. The mean facility-level missing rate monotonically increases as a function of quartile QM score, from the quartile with the lowest scores (missing rate: 10.7 percent) to the quartile with the highest scores (13.2 percent). This pattern is also shown by the significant but weak correlation between missing data and QM scores for this measure ($r = .062$, $p < .001$). Thus, facilities with higher levels of missing data rates also tend to have greater proportions of residents who self-report moderate to severe pain. This finding contradicts the hypothesis that missing data for this measure would lead to the number of residents in pain to be underreported and thus to lower (and better) scores on this measure. However, the positive relationship between missing rate and percentage of residents who self-report pain suggests that facilities with poorer performance with regard to pain management may not do well in assessing and documenting pain. Therefore, one must interpret the QM score for these facilities cautiously.

Table 4.1-9
Distribution of facility-level missing rate for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) by quartile of QM score

Quartile of QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,942	10.7%	15.9%	0.0%	0.0%	3.5%	14.6%	33.3%
26%–50%	3,840	11.5%	15.5%	0.0%	0.4%	4.9%	15.9%	33.3%
51%–75%	3,894	12.5%	16.0%	0.0%	0.8%	6.3%	17.6%	35.5%
76%–100%	3,879	13.2%	17.7%	0.0%	0.0%	5.6%	19.6%	40.0%

NOTES:

Total *n* (15,555) = number of facilities who have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced in MDS 3.0). There has been interest in the overall impact of including discharge assessments on QM rates and the completeness of the new discharge assessment items and its impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0676 by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.1-10***. The mean score for this QM is 23.1 percent when discharge assessments were included in the set of target assessments and 27.1 percent when they were excluded (mean change = 3.9 percent). Although the mean QM score only changed moderately after excluding discharge assessments from the set of eligible target assessments, the QM score change varied across facilities. The majority of facilities had an increased QM score (positive value for score change, suggesting poorer quality) after excluding discharge assessment from the set of eligible target assessments. More than 10 percent of facilities had a score change of greater than 10 percentage points. Some facilities (fewer than 25 percent) had a decreased QM score (negative value for score change, suggesting better quality) change after excluding discharge assessment from the set of eligible target assessments.

Table 4.1-10
Distribution of facility-level score change after excluding discharge assessment for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
10,957	3.9%	5.2%	-1.7%	0.3%	3.4%	7.0%	10.7%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

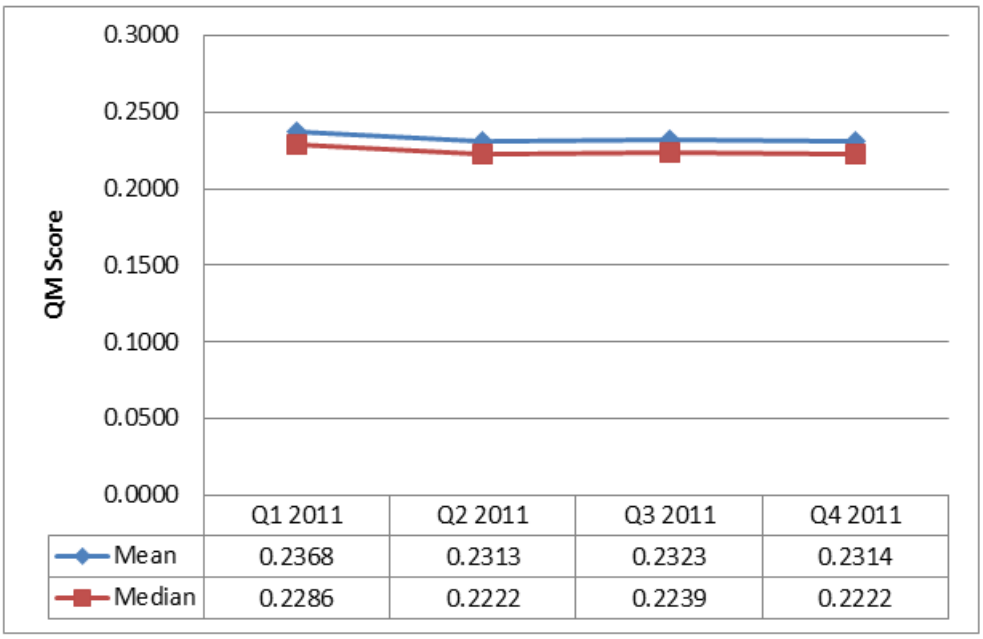
Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over time corresponding to changes in seasons, the validity of the measure is compromised, as it is being influenced by factors outside of the nursing home's control. To address whether seasonal variation might play a role in the score for QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay), we examined the national mean and median for this QM score between Quarter 1, 2011 and Quarter 4, 2011. The results are presented in ***Figure 4.1-1***.

The national-level means for this measure for each quarter between Quarter 1, 2011–Quarter 4, 2011 were 23.7 percent, 23.1 percent, 23.2 percent, and 23.1 percent. The quarterly national medians were 2.9 percent, 22.2 percent, 22.4 percent, and 22.2 percent. With 12 months of data, we did not find that the score for this QM fluctuates across quarters. However, the

absence of seasonal variation will need to be confirmed when multiple years of data become available.

Figure 4.1-1
Seasonal (quarterly) variation in QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)



SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 1–Quarter 4, 2011 \qm_quarter_1_2\complete\nh_010_10.log, \qm_quarter_2_3\complete\nh_010_10.log \qm_quarter_3_4\complete\nh_010_10.log, \qm_quarter_4_5\complete\nh_010_10.log

4.1.5 Risk Adjustment

QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay), uses sample restriction for risk adjustment, excluding residents who cannot self-report pain. This excludes a set of residents who are at risk for underreporting of pain. However, this short-stay pain prevalence measure is not risk adjusted using indirect standardization (i.e., covariate risk adjustment) the way that its counterpart, QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay), is. The short-stay population is admitted from an acute facility and represents a different case mix compared to long-stay residents. The short-stay population, particularly the post-surgical population, is likely to have acute pain which can be effectively treated and should be measured independent of risk factors. This measure was endorsed by NQF without model-based risk adjustment. There was an attempt to develop a risk adjustment model for the short-stay measure (Brega, Goodrich, Nuccio, et al., 2008). However, the explored risk adjustment models did not demonstrate adequate predictive performance.

4.2 QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

4.2.1 Summary of Findings

QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) reports the percentage of short-stay residents who had one or more new or worsening Stage 2-4 pressure ulcers during the target episode. This QM represents an improvement on prior MDS 2.0 measure because it focuses on pressure ulcers that develop within the nursing facility and excludes Stage 1 ulcers. This is an incidence measure and requires that the resident have at least one nonadmission target assessment. This QM is risk adjusted using indirect standardization, adjusting for four covariates based on the resident's initial assessment in the episode: the need for assistance in bed mobility self-performance, occasional bowel incontinence, the presence of a diabetes or peripheral vascular disease, and low BMI.

We found that 65.0 percent of the short-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, a little more than two-thirds of the facilities are able to report the risk-adjusted QM. The need for a qualifying initial assessment to obtain the covariate values had a large impact on the sample size and reportability. In the cases where no initial assessment was available an episode was not included in the QM definition. About 28 percent of the short-stay residents did not have a qualifying initial assessment. Additionally, 6.6 percent have an initial assessment but have missing values for the covariates.

In general, this was a low incidence QM (mean score: risk adjusted 1.9 percent, unadjusted 1.9 percent). Approximately one-third of facilities (35.4 percent for the risk adjusted QM, 36.2 percent for the unadjusted) had no residents with new or worsened pressure ulcers. The QM score varied across facilities with a standard deviation of 2.4 percent (2.3 percent for the unadjusted measure) and interquartile range of 2.9 percent (2.8 percent for the unadjusted measure).

To examine the reliability of this QM, we analyzed the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. About half of facilities improved by less than one standard deviation and 26.7 percent of facilities decreased in their score by less than one standard deviation. A very low proportion of facilities' scores changed over three standard deviations. About 60 percent of facilities remained within the same decile rank from quarter to quarter. In the first three quarters of 2011, about one fifth of facilities had rank changes of more than three deciles. From Quarter 3 to 4, only 6.5 percent of facilities shift more than three deciles in ranking. The relatively large rank changes in some quarters may be caused by the low incidence rate for this QM.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." It was hypothesized that changes in one QM within a measure group (e.g., short-stay and long-stay pressure ulcer QMs) should be correlated with changes in other measures because they reflect similar care processes. The findings from this analysis showed significant but weak correlations among the short-stay and long-stay pressure ulcer measures ($r = 0.148$). Second, we examined the frequency and distribution of missing data

and found that 6.6 percent of the short-stay population was excluded because they were missing data on the items used to calculate BMI. The missing rate varies across facilities, but there is little evidence to indicate a substantial relationship with the QM scores. This indicates that missing data does not pose a threat to QM validity. Finally, we analyzed the potential geographic (state) and seasonal (quarterly) variations in this QM. There was little evidence for either. The state in which the facility is located explained only about 2 percent of the variation in this QM. Regarding seasonal variation, the QM score declined only slightly from Quarter 1 to 4 in 2011.

This QM uses model-based risk adjustment of four covariates. We examined the impact of partial risk adjustment (adjusting for subsets of the four covariates) on the risk-adjusted QM scores and facility score changes. We also compared coefficients for the covariates across quarters. The results support the validity of the risk adjustment model. In addition, we explored different model specifications (single-level logistic regression models vs. hierarchical models) for risk adjustment. The results show that changing model specifications has a large impact on facility rank based on the risk-adjusted QM score.

Reliability and validity tests indicate acceptable rigor. QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) received full NQF endorsement on August 1, 2012.

4.2.2 Background and Introduction to QM

This measure captures the percentage of short-stay residents with new or worsened Stage 2–4 pressure ulcers. The measure is calculated by identifying residents with any new or worsened pressure ulcers (M0800A, M0800B, or M0800C) reported on any nonadmission assessments contained in the resident's episode. Target assessments may be discharge assessments; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or quarterly, annual, significant change, or significant correction OBRA assessments. Because this measure only captures pressure ulcers that are new since a prior assessment, residents whose target assessment is an admission assessment are excluded from this QM. This QM is risk adjusted using indirect standardization. The covariates are determined based on responses to items on the resident's initial assessment in their episode. The initial assessment is the first assessment in a resident's episode after the admission entry record (A0310F = 01). Initial assessments can be admission OBRA, 5-day PPS or discharge assessments (return not anticipated; return anticipated). Covariates used in the adjustment are: limited or more assistance in bed mobility self-performance; bowel incontinence (at least occasionally); diabetes or peripheral vascular disease; low body mass index (≥12 and ≤19). If residents only have a single assessment in their episode, or do not have complete information for the items used to construct the covariates on their initial assessment, they are excluded from the adjusted QM calculation.

A. Background for This QM

Pressure ulcers are serious medical conditions. They typically result from prolonged periods of uninterrupted pressure on the skin, soft tissue, muscle, and bone (Bates-Jensen, 2001; Institute for Healthcare Improvement, 2007; Russo, Steiner, & Spector, 2008). Vulnerable patients include the elderly; patients with stroke or diabetes, dementia, circulatory diseases,

dehydration, malnutrition, frailty, and feeding tubes; and people who use wheelchairs or are bedridden—that is, any patient with impaired mobility or sensation (Bates-Jensen, 2001; Gumieiro et al., 2012; Hurd, T., Radley, & Williams, 2010; MacLean, 2003; Teno et al., 2012). Pressure ulcers interfere with the activities of daily living, predispose patients to osteomyelitis and septicemia, and are strongly associated with longer hospital stays and mortality (Bates-Jensen, 2001).

Pressure ulcers are high-volume and high-cost adverse events across the spectrum of health care settings from acute hospitals to home health (Bates-Jensen, 2001; Hurd, T., et al., 2010; Russo et al., 2008). The prevalence of pressure ulcers in health care facilities is increasing, with some 2.5 million patients being treated annually for pressure ulcers in acute care facilities (Institute for Healthcare Improvement, 2007; Russo et al., 2008). In 2006, there were 503,300 acute hospital stays during which pressure ulcers were noted—a 78.9 percent increase from 1993 when there were about 281,300 hospital stays related to pressure ulcers (MacLean, 2003; Russo et al., 2008). Approximately 60,000 patients die from pressure ulcer complications annually (Institute for Healthcare Improvement, 2011).

Pressure ulcer incidence rates vary considerably by clinical setting—with numbers in 2007 ranging from 0.4 percent to 38 percent in acute care, from 2.2 percent to 23.9 percent in skilled nursing facilities and nursing homes, and from 0 percent to 17 percent in home care (Duncan, 2007; Institute for Healthcare Improvement, 2007). In 2009, the incidence of Stage 2 or higher pressure ulcers in nursing homes was 2.4 per 100 residents, and the prevalence was 7.3 per 100 residents (Centers for Medicare & Medicaid Services, 2010a). Incidence and prevalence are related to a facility's patient acuity, but also on their prevention practices- initiatives to improve pressure ulcer prevention can reduce incidence and prevalence, as evidenced by one four-year prevention campaign reducing the incidence from 5.19 percent to 0.73 percent (Tippet, 2009).

Patients with acute care hospitalizations related to pressure ulcers were more likely to be discharged to long-term care facilities (e.g., a skilled nursing facility, an intermediate care facility, or a nursing home), than hospitalizations for all other conditions (Hurd, T., et al., 2010; Institute for Healthcare Improvement, 2007). In fact, more than half of principal pressure ulcer stays (53.4 percent) and secondary pressure ulcer stays (54.5 percent) were discharged to long-term care—more than 3 times the rate of hospitalizations for all other conditions (16.2 percent) (Hurd, T., et al., 2010).

Pressure ulcers are one of the most important measures of the quality of clinical care in nursing facilities. The Centers for Disease Control and Prevention (CDC) conducts the National Nursing Home Survey, a continuing series of national sample surveys of nursing homes, their residents, and their staff. Data for the survey were obtained through personal interviews with facility administrators and designated staff who used administrative records to answer questions about the facilities, staff, services and programs, and medical records to answer questions about the residents. A total of 1,174 nursing home facilities participated in the latest National Nursing Home Survey (Park-Lee & Caffrey, 2009).

As reported in the 2004 National Nursing Home Survey results, about 159,000 current U.S. nursing home residents (11 percent) had pressure ulcers. Stage 2 ulcers were the most

common, accounting for about 50 percent of all pressure ulcers. Stages 1, 3, and 4 made up about the other 50 percent of all pressure ulcers (Park-Lee & Caffrey, 2009). Stage 1 pressure ulcers are not included in the proposed quality measure, because researchers have suggested that inclusion of Stage 1 pressure ulcers in the quality measures adds little value (Brega, Goodrich, Nuccio, et al., 2008). Previous research examined pressure ulcer incidence and prevalence across post-acute settings (Hurd, Moore, Radley, & Williams, 2010). For nursing homes, MDS 2.0 assessments were used for April 1, 2006, through July 15, 2006. The prevalence of pressure ulcers Stage 1-4 was 13 percent, with the prevalence of Stage 3-4 ulcers being 3 percent nationwide (Hurd, T., et al., 2010). Pressure ulcers may cause extreme discomfort to the patient and often lead to serious, life-threatening infections, which substantially increase the total cost of care (AHRQ, July 2009; National Pressure Ulcer Advisory Panel Board of Directors, 2001; Russo et al., 2008). The main driver of cost is the presence of complications, which involve diagnostic tests, additional monitoring, more expensive pressure-relieving surfaces, and extended length of stays (AHRQ, July 2009).

As reported in the Federal Register, in 2006 there were 322,946 reported cases of Medicare patients with a pressure ulcer as a secondary diagnosis—each case had an average charge of $40,381 for a hospital stay, for an annual total cost of $13 billion (Centers for Medicare & Medicaid Services, 2007). In 2008, there were 394,699 cases of pressure ulcers in Medicare patients, with the cost of treating the pressure ulcer alone at $8,730 per case, resulting in a total cost of over $3.4 billion for the pressure ulcer treatment (Van Den Bos et al., 2011). To address this critical clinical issue, there are numerous national health care organizations with ongoing efforts and publications to prevent pressure ulcers, monitor prevalence, and improve treatment. One national campaign, The Advancing Excellence in America's Nursing Homes Campaign, was launched in 2006 to help nursing homes measurably to improve care in several clinical areas including pressure ulcers. Results from Phase 2 of the campaign show that campaign nursing homes selecting the goal of preventing pressure ulcers reduced the prevalence of high-risk pressure ulcers at a faster pace than other nursing homes in 2010 (Colorado Foundation for Medical Care, 2011). In addition, the campaign has documented that it can cost as much as $19,000 to treat a single Stage 4 pressure ulcer.

B. Summary of Differences between MDS 2.0 to MDS 3.0 QM Definitions

The MDS 2.0 set of items used to construct this measure were revised with implementation of the MDS 3.0. As Saliba and Buchanan (2008) noted during the development of the MDS 3.0, whenever possible, they included items or language used in other health care settings in order to improve communication across settings and providers (e.g., the pressure ulcer items included in the National Pressure Ulcer Advisory Panel's Pressure Ulcer Scale for Healing [PUSH] tool are used to describe pressure ulcers in the MDS 3.0). Therefore, the proposed measure based on the new MDS 3.0 pressure ulcer items better aligns the measure with accepted best practices.

QM item-level enhancement on the MDS 3.0 includes:

- eliminating reverse staging, which does not reflect the pathophysiology of pressure ulcer healing;

- staging pressure ulcers based on deepest anatomical change (recommended by Wound, Ostomy, and Continence Nurses Society [WOCN], National Pressure Ulcer Advisory Panel [NPUAP]);
- assessing unstageable pressure ulcers as separate items (recommended by NPUAP, WOCN);
- collecting the number of pressure ulcers that were present on admission for each stage;
- determining the tissue type for pressure ulcers in the most advanced stage;
- using definitions based on NPUAP; and
- assessing unstageable pressure ulcers as separate items (recommended by NPUAP, WOCN).

Stage 1 pressure ulcers are not included in the MDS 3.0 pressure ulcer quality measure. Researchers indicated that including Stage 1 pressure ulcers in quality measures adds little value, penalizes facilities for early identification, and that Stage 1 pressure ulcers are inconsistently assessed, especially for populations with darker skins. Another design refinement is respecifying this measure as an incidence measure, holding facilities accountable only for those pressure ulcers that developed or worsened while the facility was providing services to the resident.

Note also that the resident population selection changed from MDS 2.0 to MDS 3.0. In the MDS 2.0 measure, residents were included as post-acute care if they had a 14-day PPS MDS in the target quarters; the measure was calculated based on just 14-day PPS assessments. In the MDS 3.0 measure, the analogous short-stay population comprises residents with 100 or fewer cumulative days in facility and whose QM may be calculated based on discharge, 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments, or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections present findings regarding QM #0678 using MDS 3.0 episode file for Quarter 4, 2011(unless otherwise specified). Four general areas were analyzed:

- Number of assessments included and excluded from the numerator of the QM based on sample restrictions.
- Whole sample prevalence of items that are related to the QM
- Findings regarding variability, reportability, reliability and validity
- Risk adjustment analyses.

4.2.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 1,325,274 (96.0 percent) assessments included in the denominator of QM #0678 Percent of Residents With Pressure Ulcers That Are New or

Worsened (Short Stay), and a total of 55,587 resident episodes (4.0 percent of all short-stay resident episodes) were excluded from the denominator. ***Table 4.2-1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criteria. Note that one resident episode can meet multiple exclusion criteria. The number and percent of resident episodes that met each exclusion criterion are as follows:

- A total of 55,574 (4.0 percent) assessments met exclusion criterion 1 (missing data).
- Thirteen assessments (0.0 percent) met exclusion criterion 2 (inconsistent: none of the three items M0800A, M0800B, and M0800C is usable).
- Most resident episodes (389,651, or 28.2 percent) met exclusion criterion 3 and were excluded due to not having a paired initial assessment needed to calculate the measure; and
- 90,713 resident episodes (6.6 percent) met exclusion criterion 4 (missing covariate data on items used to calculate BMI)

Table 4.2-1
Target assessments used in the calculation of QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	897,206	65.0%
Excluded—Exclusion 1	55,754	4.0
Excluded—Exclusion 2	13	0.0
Excluded—Exclusion 3	389,651	28.2
Excluded—Exclusion 4	90,713	6.6
Total number of short-stay resident episodes	1,381,028	—

NOTES:

Exclusion 1 = Missing data

Exclusion 2 = Inconsistent: none of the three items M0800A (Stage 2 Pressure Ulcers), M0800B (Stage 3 Pressure Ulcers), and M0800C (Stage 4 Pressure Ulcers) is usable.

Exclusion 3 = Missing initial assessment

Exclusion 4= Missing covariate data on items used to calculate BMI

[a] Column values may not add up to total since a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident assessments in short-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

In this section, we present a series of descriptive tables regarding residents included in the QM. ***Table 4.2-2*** shows frequencies of new and worsened pressure ulcers since the last assessment, by stage. Note that this table underestimates the total set of episodes with any

occurrence of new and worsened pressure ulcers captured by the QM numerator because it only captures information from one point in time (the target assessment) rather than the full set of assessments included in a resident's episode. The most common stage of new or worsened ulcer reported below is Stage 2 (4.7 percent), followed by Stage 3 (1.4 percent) and Stage 4 (0.9 percent).

Table 4.2-2
Frequencies of new and worsened pressure ulcers since prior assessment as reported on target assessments in QM #0678
Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay), by stage, covariate exclusions not applied

Count of pressure ulcers	M0800A Worsening or New Stage 2 Pressure Ulcer *n*	M0800A Worsening or New Stage 2 Pressure Ulcer %	M0800B Worsening or New Stage 3 Pressure Ulcer *n*	M0800B Worsening or New Stage 3 Pressure Ulcer %	M0800C Worsening or New Stage 4 Pressure Ulcer *n*	M0800C Worsening or New Stage 4 Pressure Ulcer %
No response*	4,232	0.3%	4,294	0.32%	4,281	0.32%
0	73,128	5.5%	118,196	8.92%	125,089	9.44%
1	46,099	3.5%	14,706	1.11%	8,998	0.68%
2	12,272	0.9%	2,477	0.19%	1,459	0.11%
3	3,268	0.2%	722	0.05%	523	0.04%
4	990	0.1%	193	0.01%	216	0.02%
5	373	0.0%	66	0.00%	76	0.01%
6	164	0.0%	27	0.00%	34	0.00%
7	80	0.0%	16	0.00%	14	0.00%
8	45	0.0%	9	0.00%	8	0.00%
9	68	0.0%	13	0.00%	21	0.00%
No unhealed pressure ulcer	1,184,555	89.4%	1,184,555	89.4%	1,184,555	89.4%
Total	1,325,274	100.0%	1,325,274	100.0%	1,325,274	100.0%

*Assessments with at least one valid response indicating a new or worsened ulcer retained in the QM definition.

Analysis date: 6/28/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db166_request\db166_request.log)

Table 4.2-3 shows frequencies on pressure ulcers by stage on the target assessment in the episode. We present the frequencies separately for residents counted in the numerator and those not counted in the numerator. This is a snapshot of the burden of pressure ulcers in the total population at one point in time (disregarding whether or not a pressure ulcer was new or had worsened). Table 4.2-3 shows similar results as Table 4.2-2 (i.e., it shows that the most common pressure ulcer stage is Stage 2 across both populations—residents who trigger the QM with new or worsened ulcers [60.9 percent] and residents who do not trigger the QM [3.9 percent]). Not surprisingly, a much larger proportion of residents with new and worsened pressure ulcers during their episodes had more pressure ulcers at the time of their target assessment (e.g., 60.9 percent had at least one Stage 2 compared to just 3.9 percent of nontriggering residents). It is worth noting that while the proportion of pressure ulcers is very low among residents who did not trigger this QM, these residents represent the majority of all residents who have any Stage 2 ulcers (over 50,267 versus 13,092 residents who trigger the QM). However, when pressure ulcers present on admission are separated out (***Table 4.2-4***), it becomes clear that the majority of the nontriggering residents had ulcers that were present on admission (42,687 of 50,267, compared to 5,702 of 13,092).

4.2.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with "perfect scores" (i.e., 0 percent rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes) indicating a "ceiling effect." ***Table 4.2-5*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) is 1.9 percent with a standard deviation of 2.4 percent. This QM has an interquartile range of 2.9 percent. About 35.4 percent of facilities have scores of 0 percent ("perfect" scores).

Table 4.2-3
Frequencies by stage of pressure ulcers, by whether target assessment is in QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) numerator, by stage, covariate exclusions not applied

Count of pressure ulcers (PU)	Stage 2 PU doesn't trigger QM *n*	Stage 2 PU doesn't trigger QM %	Stage 2 PU triggers QM *n*	Stage 2 PU triggers QM %	Stage 2 PU total %	Stage 3 PU doesn't trigger QM *n*	Stage 3 PU doesn't trigger QM %	Stage 3 PU triggers QM *n*	Stage 3 PU triggers QM %	Stage 3 PU total %	Stage 4 PU doesn't trigger QM *n*	Stage 4 PU doesn't trigger QM %	Stage 4 PU triggers QM *n*	Stage 4 PU triggers QM %	Stage 4 PU total %
No response	4,127	0.3%	105	0.5%	0.3%	4,166	0.3%	128	0.6%	0.3%	4,154	0.3%	127	0.0%	0.6%
0 of this stage	68,488	5.3	4,640	21.6	5.5	104,142	8.0	14,054	65.4	8.9	109,311	8.4	15,778	1.2	73.4
At least 1 of this stage	50,267	3.9	13,092	60.9	4.8	14,574	1.1	3,655	17.0	1.4	9,417	0.7	1,932	0.1	9.0
No unhealed	1,180,894	90.0%	3,661	17.0	89.4	1,180,894	90.6	3,661	17.0	89.4	1,180,894	90.6	3,661	0.3	17.0
Total	1,303,776	100.0	21,498	100.0	100.0	1,303,776	100.0	21,498	100.0	100.0%	1,303,776	100.0	21,498	100.0	100.0

Analysis date: 6/28/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db166_request\db166_request.log)

Table 4.2-4

Frequencies of pressure ulcers that were present on admission, by whether the target assessment is in QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) numerator, by stage, covariate exclusions not applied

Count of pressure ulcers	Stage 2 PU doesn't trigger QM *n*	Stage 2 PU doesn't trigger QM %	Stage 2 PU triggers QM *n*	Stage 2 PU triggers QM %	Stage 2 PU total %	Stage 3 PU doesn't trigger QM *n*	Stage 3 PU doesn't trigger QM %	Stage 3 PU triggers QM *n*	Stage 3 PU triggers QM %	Stage 3 PU total %	Stage 4 PU doesn't trigger QM *n*	Stage 4 PU doesn't trigger QM %	Stage 4 PU triggers QM *n*	Stage 4 PU triggers QM %	Stage 4 PU total %
No response	4,186	0.3%	128	0.0%	0.6%	4,183	0.3%	141	0.7%	0.3%	4,163	0.3%	134	0.6%	0.3%
0 of this stage	7,521	0.6	7,367	0.6	34.3	1,400	0.1	1,582	7.4	0.2	446	0.0%	431	2.0	0.1
At least 1 of this stage	42,687	3.3	5,702	0.4	26.5	13,157	1.0	2,060	9.6	1.1	8,962	0.7%	1,494	6.9	0.8
No unhealed	1,249,382	95.8	8,301	0.6	38.6	1,285,036	98.6	17,715	82.4	98.3	1,290,205	99.0%	19,439	90.4	98.8
Total	1,303,776	100.0	21,498	1.6	100.0	1,303,776	100.0	21,498	100.0	100.0	1,303,776	100.0	21,498	100.0	100.0

Analysis date: 6/28/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db166_request\db166_request.log)

Table 4.2-5
QM score distribution for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

Risk adjustment used	*n*	Mean score	Std dev.	10th percen-tile	25th percen-tile	50th percen-tile	75th percen-tile	90th percen-tile	% of facilities with "perfect scores"	Inter-quartile range
Unadjusted	12,631	1.9%	2.3%	0%	0%	1.2%	2.8%	4.8%	36.2%	2.8%
Adjusted	10,763	1.9%	2.4%	0%	0%	1.3%	2.9%	4.9%	35.4%	2.9%

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\combined_fa_adj\nh_012_10_ra_combined_v4.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 20 residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examine the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.2-6*** shows the results using Quarter 4, 2011. After applying measure exclusion criteria, 68.6 percent of facilities had sufficient sample size to meet minimum requirements for public reporting on the risk-adjusted measure. This is an improvement over the MDS 2.0 measure, for which 65.1 percent of facilities were able to report (Quarter 1, 2006) (Brega, Goodrich, Hittle, Conway, & Levy, 2008).

Table 4.2-6
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator [a]	897,206 (65.0%)[a]	538,146 (39.0%)[a]
Number (%) of resident episodes in the numerator	21,498 (1.6%)[a]	14,713 (1.1%)[a]
Number (%) of facilities able to report this QM[b]	10,763 (68.6%)[b]	8,311 (53.0%)[b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the short-stay sample (1,381,028).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay residents (15,686).

Analysis date: 5/29/2012

SOURCE: MDS 3.0, RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log, nh_012_10_nd and nh_012_10.log)

The inclusion of discharge assessments is an important change from the MDS 2.0 to the MDS 3.0 measure. Including the discharge assessment may potentially increase the number of resident episodes included in QM calculations. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 specifications. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30- day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay population for Quarter 4, 2011 were discharge assessments, and that 8.8 percent of target assessments for the long-stay population in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may affect the reportability of short-stay QMs to a greater degree than the long-stay QMs.

To evaluate the impact on reportability of including discharge assessments in the set of target assessments eligible for constructing QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay), we report changes in the number of facilities able to report this QM when discharge assessments are not included in the set of target assessments. The results are presented in Table 4.2-6. After excluding discharge assessments from the set of target assessments, reportability substantially decreased for QM #0678 (68.6 percent with discharge assessments vs. 53.0 percent without discharge assessments).

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from quarter 3 to quarter 4 of 2011.

Table 4.2-7 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. For each quarter-to-quarter transition during the year 2011, the majority of facility rankings remained within the same decile. However, relatively large numbers of facilities saw large changes (>3 deciles) in the ranking. The relatively large rank changes in some quarters may be caused by the low incidence rate for this QM.

Table 4.2-7
Distribution of change in facility rank from one quarter to the next for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	5,962 (59.9%)	6,090 (59.1%)	6,257 (60.6%)
Between 1 and 2 deciles	1,204 (12.1%)	1,117 (10.9%)	1,091 (10.6%)
Between 2 and 3 deciles	971 (9.8%)	868 (8.4%)	726 (7.0%)
More than 3 deciles	1,813 (18.2%)	2,224 (21.6%)	2,247 (21.8%)
Total	9,950	10,299	10,321

NOTES:

Total number of facilities in each column reflects all the facilities that could report the measure in both quarters.

Analysis date: 6/15/2012

SOURCE: MDS 3.0, RTI analysis of MDS 3.0 episode files for Quarter 1–Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.2-8 presents the results from an analysis of the change in facility scores for this quality measure between quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 2.4 percent). Overall, the average change in scores for this measure was a decline of 0.2 percent. The vast majority of facilities had very little change in their score from quarter to quarter (78.6 percent facilities changed less than one standard deviation), whereas less than 2 percent facilities changed over three standard deviations.

Table 4.2-8
Change in facility score from one quarter to the next for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
10,321	-0.2%	2.1%	1.0%	2.2%	7.9%	26.7%	51.9%	7.7%	1.9%	0.8%

NOTES:

Number of facilities reporting (10,321) reflects facilities that reported this QM in both Quarter 3 and Quarter 4, 2011.

Analysis Date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated). For example, one could argue that the "pressure ulcer group" of measures, QMs #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) and #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay), should reflect the quality of related clinical care processes. Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) and the facility's percentile rank on QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) for Quarter 4, 2011. We found that the correlation between these two QMs was weak ($r = 0.1482$) but statistically significant ($p < 0.001$).[8] The weak correlation may be due to the different specifications for the short-stay and long-stay QMs (i.e., the short-stay QM is an incidence measure, while the long-stay QM is a prevalence measure). It may also suggest that the care processes for preventing pressure ulcers and/or the patient characteristics in the short-stay versus long-stay populations are very different and that facilities may have more specialized expertise in dealing with only one of the patient groups.

Variation by State. For a quality measure to be valid, variation observed in the distribution of the QM should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as state-level payment policies or demographics, such variation can be a threat to the validity of the measure. To explore whether state characteristics might be a source of facility score variation for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) and thus a potential threat to the measure's validity we conducted ANOVA. We also examined the interquartile difference between the mean state-level scores across states. The proportion of variance in this measure explained by the state in which facilities are located is 2.0 percent and significant [$F(52, 10{,}710) = 4.3$, $p < .001$].[9] The difference between the mean state-level scores for states at the 25th percentile and the 75th percentile is 0.5 percentage point. Thus, there is very little variance explained by geographic location for this quality measure. However, the proportion is significant, and this issue merits future monitoring.

---

[8] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

[9] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns indicate that certain types of residents tend to have assessments with missing data in ways that affect the calculation of a quality measure, then that measure may not be capturing processes and outcomes for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

In Quarter 4, 2011, 483,822 short-stay residents (35.0 percent) were excluded from the calculation of this measure. Of these, 55,754 (4.0 percent of short-stay residents) were excluded because of missing or invalid responses on items indicating the presence of new or worsening Stage 2-4 pressure ulcers on all assessments in the look-back period. A small number of exclusions (13, or less than one tenth of 1 percent of the short-stay population) were excluded because these items were unusable. The majority of exclusions—389,651 (28.2 percent of the short-stay population)—were due to missing initial assessments. Finally, 90,713 residents (6.6 percent of the short-stay population) were excluded because they were missing data on the items used to calculate BMI ([0] values are imputed for missing data on other items involved in covariate calculations).

Table 4.2-9 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. Facility-level missing data rates for this measure are low, with a mean of 5.3 percent. However, the calculation of this measure does allow for imputation of values for missing covariate data (except for BMI). Facility rates appear to be distributed fairly evenly, indicating that there is likely no subset of facilities that is disproportionately affected.

We further examined the possible relationship between missing data and QM scores. ***Table 4.2-10*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure. There appears to be no linear relationship between mean facility-level missing-data rates and QM scores. For this measure, there is a weak positive correlation between missing data and QM scores ($r = 0.057$, $p < .001$). Thus, facilities with more missing data tend to have slightly higher scores for this measure. Since missing-data rates predict less than one-half of 1 percent of the variance in QM scores for this measure, they likely do not present a threat to the validity of this measure.

Table 4.2-9
Distribution of facility-level missing rate for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,631	5.3%	6.3%	0%	1.2%	3.8%	7.3%	12%

NOTES:

n = number of facilities who have data for numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request_012_new.log)

Table 4.2-10
Distribution of facility-level missing rate for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

Quartile of QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	6,991	5.2%	7.2%	0.0%	0.0%	3.3%	7.3%	12.7%
26%–50%	831	5.0%	4.8%	1.2%	2.2%	3.7%	6.4%	9.8%
51%–75%	3,954	5.1%	4.5%	1.0%	2.2%	4.1%	6.8%	10.4%
76%–100%	3,851	5.6%	5.7%	0.0%	1.7%	4.3%	7.9%	12.5%

NOTES:

Total *n* (15,627) = number of facilities who have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 Data episode file for Quarter 4, 2011 (\quarter_4_5\db181_request_012_new \db181_request_012_new.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in the overall impact of including the discharge assessment on QM rates and the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluate the impact of the new discharge assessment on QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.2-11***. The mean nursing home level QM score and distributions change slightly before and after excluding discharge assessments from the set of eligible target assessments (mean change = -0.1 percent). The vast majority of facilities had small score change after excluding discharge assessments from the set of eligible target assessments.

Table 4.2-11
Distribution of facility-level score change after excluding discharge assessment for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
8,311	-0.1%	1.4%	-1.7%	-0.4%	0.0%	0.5%	1.2%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

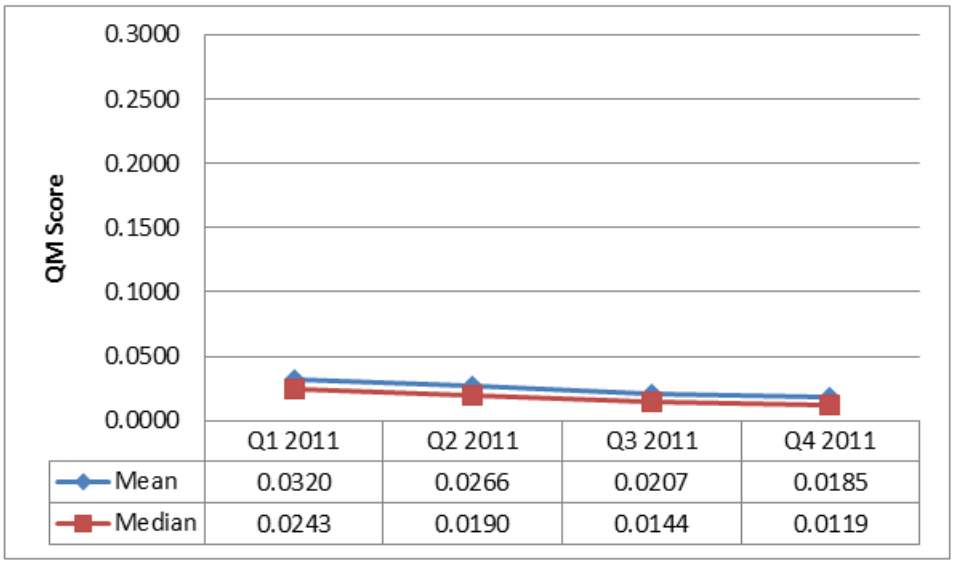
Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 Data episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because the QM score is being influenced by factors outside of the nursing home's control. NQF expressed concern that QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) could be subject to seasonal variation. To address this concern, we examined the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.2-2***.

Figure 4.2-2
Seasonal (quarterly) variation in QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)



Analysis date:

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 1–Quarter 4, 2011 \qm_quarter_1_2\complete\nh_012_10.log \qm_quarter_2_3\complete\nh_012_10.log \qm_quarter_3_4\complete\nh_012_10.log; \qm_quarter_4_5\complete\nh_012_10.log

The national-level means for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, were 3.2 percent, 2.7 percent, 2.1 percent, and 1.9 percent. The quarterly national medians were 2.4 percent, 1.9 percent, 1.4 percent, and 1.1 percent. Both mean and median consistently decreased over the 4 quarters. With only 12 months of data available, it is too early to ascertain whether the changes in national QM score over time reflect seasonal variation; this will need to be investigated further when multiple years of data become available.

4.2.5 Risk Adjustment

QM#0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) is model-based, risk-adjusted with four covariates defined based on responses on residents' initial assessment in their episode, including: requires assistance with bed mobility; bowel

incontinence at least once a week; presence of diabetes or peripheral vascular disease; and low body mass index (BMI equal to or greater than 12 and less than or equal to 19).

Covariates for this measure are based on the initial assessment in the episode, which is the first assessment following the admission entry record (A0310F = 01); assessments can be admission, a 5-day PPS assessment, or a discharge assessment (return not anticipated and return anticipated). More details about the methods for calculating model-based risk-adjusted measures are included in Section 4 of this document and in Appendix B, which includes the MDS 3.0 QM User's Manual.

The frequencies of the covariates used in the short-stay pressure ulcer quality measure among residents included in the QM definition are shown in ***Table 4.2-12***. Note that these analyses contain only episodes with valid initial assessments, but that the QM definition excludes episodes with any missing covariate values, which were most frequent with diabetes/peripheral vascular disease (PVD). There were no missing values for bed mobility or bowel incontinence on the initial assessments identified. The majority of residents (89.2 percent) in the adjusted QM have impairments in bed mobility self-performance. Diabetes and PVD were present in less than half of the episodes and bowel incontinence in slightly more than one third. Low BMI was present in about 10 percent of cases.

Facility QM scores may be influenced by combinations of covariates. Analyses (***Table 4.2-13***) were performed to explore the mean adjusted QM scores using different combinations of the four covariates used to calculate this QM. These analysis also provide the number and percent of facilities where the combination of covariates used would cause the facility's score to be different from the fully risk adjusted score. When all four covariates are used the average percent of short-stay residents with new or worsened pressure ulcers is 2.5 percent, this is the fully risk adjusted score. When only the indicator for assistance in bed mobility is used, the average percent of short-stay residents with new or worsened pressure ulcers remains the same, but 52 percent of facilities would have QM scores that are lower than the fully risk adjusted score, and the remaining 48 percent would have scores that are higher than the fully adjusted score. The combination of covariates that have the greatest negative impact on facilities' scores are those risk adjusted by all covariates *except* the bed mobility assistance item. Applying this combination of covariates, 76 percent of facilities have a higher percent of short-stay residents with new or worsened pressure ulcers (a lower or worse QM score).

Table 4.2-12
Frequencies on covariate values for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

Covariate	Episodes with valid values (*n*)	Percent with covariate condition
Bed mobility: require limited or more assistance	991,377	89.2%
Bowel incontinence (occasional or more)	991,377	34.5%
Diabetes or peripheral vascular disease	859,666	42.4%
Low body-mass index	900,664	9.8%

Analysis date: 4/17/2012

SOURCE: RTI Analysis of MDS 3.0 episode file for Quarter 4, 2011 (qm_quarter_4_5\complete\nh_012_10_ra.log)

Table 4.2-13
Impact of varying combination of covariates in risk adjustment for this measure frequency of covariate combinations used for risk adjustment: QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

Cov. 1	Cov. 2	Cov. 3	Cov. 4	Mean	SD	p50	*n*	Facilities whose QM increases by 5% of QM mean when covariate(s) omitted *n*	Facilities whose QM increase by 5% of QM mean when covariate(s) omitted %	Facilities whose QM decreases by 5% of QM mean when covariate(s) omitted *n*	Facilities whose QM decreases by 5% of QM mean when covariate(s) omitted %
1	1	1	1	2.5%	3.6%	1.4%	7,214	—	—	—	—
1	0	0	0	2.5%	3.4%	1.5%	9,150	1,584	17.3%	4,490	49.1%
1	1	0	0	2.5%	3.3%	1.5%	9,150	1,210	13.2%	5,007	54.7%
1	1	1	0	2.5%	3.4%	1.4%	8,098	810	10.0%	3,467	42.8%
1	0	1	0	2.5%	3.5%	1.4%	8,098	1,525	18.8%	3,080	38.0%
1	0	1	1	2.5%	3.7%	1.3%	7,214	1,659	23.0%	1,594	22.1%
1	0	0	1	2.4%	3.5%	1.4%	8,180	1,507	18.4%	3,177	38.8%
0	1	0	0	2.4%	3.1%	1.5%	9,150	1,217	13.3%	4,910	53.7%
0	1	1	0	2.4%	3.3%	1.4%	8,098	880	10.9%	3,222	39.8%
0	1	1	1	2.5%	3.4%	1.4%	7,214	1,000	13.9%	1,067	14.8%
0	1	0	1	2.4%	3.2%	1.4%	8,180	821	10.0%	3,486	42.6%
0	0	1	0	2.4%	3.3%	1.4%	8,098	1,540	19.0%	3,094	38.2%
0	0	1	1	2.4%	3.4%	1.4%	7,214	1,653	22.9%	1,688	23.4%
0	0	0	1	2.4%	3.3%	1.4%	8,180	1,493	18.3%	3,199	39.1%
0	0	0	0	2.5%	3.6%	1.2%	12,630	1,688	13.4%	1,742	13.8%

NOTES:

Total number of facilities with sample size ≥ 20 is 12,630

Covariate1—assistance in bed mobility

Covariate2—any bowel incontinence

Covariate3—diabetes or peripheral vascular disease

Covariate4—low body mass index

Analysis date: 6/28/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db169_request\db169_request.log)

This analysis illustrates that the selection and combination of covariates used for risk adjustment greatly influences facilities' QM scores and underscores the importance of parsimonious selection and expert clinical guidance in model specification.

To examine whether the performance of the risk-adjustment model for this QM is stable across quarters, we compared the intercepts and coefficients for the covariates for each quarter in 2011. In ***Table 4.2-14***, we show that values for the intercept are very similar across quarters. Coefficients for the covariates also do not change much from Quarter 1 to Quarter 4, 2011, except that the coefficient for diabetes or PVD dropped from 0.430 in Quarter 1 to 0.367 in Quarter 2. This coefficient continued to slightly decrease in Quarter 3 and 4. The model goodness-of-fit statistic—pseudo R-square—also has comparable values across four quarters, ranging from 0.032 to 0.037. Overall, these results suggest that the performance of the risk adjustment is stable in 2011 and thus support the validity of the risk-adjustment model.

To understand how model specification influences the performance of the risk-adjustment model and the adjusted QM scores, we estimated the risk adjustment model for this QM using both a single-level model and a hierarchical model. The current specification of the risk adjustment model for this QM uses single-level logistic regression, which does not take into account the fact that residents are clustered within nursing homes. The hierarchical model, however, addressed the issue. We estimated the hierarchical model using the specification of multilevel mixed-effects logistic regression. ***Table 4.2-15*** shows that coefficients for the covariates from the single-level model and those from the hierarchical model are very similar, but the hierarchical model has a much higher predictive power. The C-statistics indicate that the single-level model has acceptable power for predicting pressure ulcers that are new or worsened (C-stat = 0.654), and the hierarchical model has nearly excellent predictive power for this QM (C-stat= 0.791).

The risk adjusted QM score based on single-level model is different from those based on the hierarchical model. For an average facility, the risk-adjusted QM score based on the single-level model is slightly higher than the one based on the hierarchical model (difference = 0.6 percent). The interquartile range for facility-level difference in risk adjusted QM score between single-level and hierarchical models is -0.1 percent to 0.4 percent. Based on these risk adjusted QM scores, we also compared facility rank, with a particular interest in facilities between 95 and 100 percentile (indicating extremely poor quality of care). Among the 444 facilities identified by the single-level model as poor-quality outliners, only 257 (57.9 percent) were identified as poor-quality outliners based on the hierarchical model. [10]

These results suggest that the risk-adjustment model for this QM performs differently with different model specifications. As a result, the risk-adjusted QM scores and facility rank based on these scores show differences.

---

[10] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db130_request\db130_request.log)

Table 4.2-14
Intercepts and coefficients by quarter of data for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)

Covariate in 2011	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Intercept	-5.170	-5.225	-5.340	-5.354
Bed mobility	1.028	1.026	1.021	0.992
Bowel incontinence	0.866	0.860	0.858	0.860
Diabetes or peripheral vascular disease	0.430	0.367	0.359	0.350
Low body mass index	0.402	0.402	0.391	0.398
Pseudo R-square	0.037	0.034	0.033	0.032

NOTES:

1. Indicator of requiring limited or more assistance in bed mobility self-performance dependence on the initial assessment:

Covariate = [1] if G0110A1 = [2, 3, 4, 7, 8]

Covariate = [0] if G0110A1 = [0, 1, −]

2. Indicator of bowel incontinence at least occasionally on the initial assessment:

Covariate = [1] if H0400 = [1, 2, 3]

Covariate = [0] if H0400 = [0, 9, − , ^]

3. Have diabetes or peripheral vascular disease on initial assessment:

Covariate = [1] if any of the following are true:

a. I0900 = [1] (checked)

b. I2900 = [1] (checked)

c. I8000A through I8000J contains any of the following peripheral vascular disease diagnosis codes: [250.7, 440.20, 440.21, 440.22, 440.23, 440.24, 440.29, 440.31, 440.32, 443.81, 443.9].

Covariate = [0] if I0900 = [0, −] AND I2900 = [0, - -]AND I8000A through I8000J do not contain any of the peripheral vascular disease diagnosis codes listed above.

4. Indicator of low body mass index, based on height (K0200A) and weight (K0200B) on the initial assessment:

Covariate = [1] if BMI ≥ [12.0] AND ≤ [19.0]

Covariate = [0] if BMI > [19.0] AND ≤ [40.0]

Where: BMI = (weight * 703 / height2) = ((K0200B) * 703) / (K0200A2) and the resulting value is rounded to one decimal.

Covariate = missing if K0200A = [−] OR K0200B = [−] OR BMI < [12.0] OR BMI > [40.0].

Analysis date: 4/17/2012

SOURCE: RTI Analysis of MDS 3.0 episode files for Quarter 1 – Quarter 4, 2011(\qm_quarter_4_5\complete\nh_012_10_ra.log; qm_quarter_3_4\complete\nh_012_10_ra.log; qm_quarter_2_3\complete\nh_012_10_ra.log; qm_quarter_1_2\complete\nh_012_10_ra.log)

Table 4.2-15
Intercepts and coefficients for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay)—Comparison between single-level model and hierarchical generalized linear model

Covariate in Quarter 4, 2011	Single-level model covariate	Single-level model 95% confidence interval	Hierarchical model covariate	95% hierarchical model confidence interval
Bed mobility	0.992	0.903–1.080	1.079	0.989–1.169
Bowel incontinence	0.860	0.828–0.892	0.878	0.845–0.911
Diabetes or peripheral vascular disease	0.350	0.319–0.381	0.354	0.322–0.386
Low body mass index	0.398	0.353–0.444	0.405	0.359–0.452
C-statistic	0.654	—	0.791	—

NOTES:

1. Indicator of requiring limited or more assistance in bed mobility self-performance dependence on the initial assessment:

Covariate = [1] if G0110A1 = [2, 3, 4, 7, 8]

Covariate = [0] if G0110A1 = [0, 1, −]

2. Indicator of bowel incontinence at least occasionally on the initial assessment:

Covariate = [1] if H0400 = [1, 2, 3]

Covariate = [0] if H0400 = [0, 9, − , ^]

3. Have diabetes or peripheral vascular disease on initial assessment:

Covariate = [1] if any of the following are true:

a. I0900 = [1] (checked)

b. I2900 = [1] (checked)

c. I8000A through I8000J contains any of the following peripheral vascular disease diagnosis codes: [250.7, 440.20, 440.21, 440.22, 440.23, 440.24, 440.29, 440.31, 440.32, 443.81, 443.9]

Covariate = [0] if I0900 = [0, −] AND I2900 = [0, - -]AND I8000A through I8000J do not contain any of the peripheral vascular disease diagnosis codes listed above

4. Indicator of Low Body Mass Index, based on Height (K0200A) and Weight (K0200B) on the initial assessment:

Covariate = [1] if BMI ≥ [12.0] AND ≤ [19.0]

Covariate = [0] if BMI > [19.0] AND ≤ [40.0]

Where: BMI = (weight * 703 / height2) = ((K0200B) * 703) / (K0200A2) and the resulting value is rounded to one decimal

Covariate = missing if K0200A = [−] OR K0200B = [−] OR BMI < [12.0] OR BMI > [40.0]

Analysis date: 6/4/2012

SOURCE: RTI Analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db130_request\nh_012_10_xtra_state.log)

4.3 QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Influenza Vaccination (Short Stay)

4.3.1 Summary of Findings

QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) reports the percentage of short-stay residents who either received the influenza vaccine during the current or most recent influenza season (either in the facility or outside the facility) or were offered and declined the vaccine or were ineligible due to contraindication(s). This is an important measure of quality of care in the nursing facility, as morbidity and mortality related to influenza are often reported in conjunction with data regarding pneumonia, and together frequently lead to death in the elderly population.

About 71.7 percent of the short-stay resident episodes meet the denominator inclusion criteria and were included in the denominator. Based on these resident episodes, 75.4 percent of facilities have 20 or more short-stay episodes included in the denominator and are able to report this QM. Among facilities able to report, the mean facility QM score is 79.7 percent. The QM score varies across facilities with a standard deviation of 19.2 percent and an interquartile range of 23.3 percent. This indicates acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To examine the reliability of this QM, changes in QM score and in rank based on the QM score for each facility were examined by quarter. About half of facilities declined in their score by less than one standard deviation and 26.7 percent of facilities decreased in their score by less than one standard deviation. In a very low proportion of facilities score changes were greater than three standard deviations. A similar pattern was found for facility percentile ranking changes. About two-thirds of facilities remained within the same decile from quarter to quarter. Almost all facilities had rank changes within 3 deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group (e.g., vaccination QMs) should be correlated with changes in other measures because they reflect similar care processes. Findings from this analysis showed moderate to high correlations among the vaccination QMs. Given that values of zero are imputed for missing data on items used to calculate this quality measure, no resident episode was excluded due to missing data. We examined the imputation rate across facilities and found very weak relationship between imputation rate and facility QM score. The result suggests that missing data should not be a threat to the validity of this QM. We also analyzed the potential geographic (state) and seasonal (quarterly) variations in this QM. The state in which the facility is located explains about 4 percent of the variation in this QM. The QM score slightly fluctuates across quarters in 2011 and peaks in Quarter 2, 2011.

4.3.2 Background and Introduction to QM

The quality measure reports the percent of short-stay residents who are appropriately given the influenza vaccination during the current or most recent influenza season. The measure is reported as the aggregate of three separately calculated submeasures to reflect the various

ways that a resident may be "appropriately" given the influenza vaccination during the current or most recent influenza season. The three submeasures are also calculated and reported separately to harmonize with the NQF and CDC reporting structure. The three submeasures are:

- resident received the influenza vaccine during the current or most recent influenza season, either in the facility or outside the facility;
- resident was offered and declined the influenza vaccine; and
- resident was ineligible to receive the influenza vaccine due to contraindication(s) (e.g., anaphylactic hypersensitivity to eggs or other components of the vaccine, history of Guillain-Barré syndrome within 6 weeks after a previous influenza vaccination, bone marrow transplant within the past 6 months).

A. Background of QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

Morbidity and mortality data related to influenza are often reported in conjunction with data regarding pneumonia. According to CDC, pneumonia and influenza were the seventh most common cause of death for persons aged 65 and older in the United States (Centers for Disease Control and Prevention, 2009). In 2009, influenza and pneumonia combined caused 43,465 deaths in people over the age of 65, with 638 deaths caused by influenza alone. In addition to being at risk for primary illness, frail elderly are especially vulnerable and subject to complications of influenza. In 2004, there were approximately 123,000 deaths with influenza and pneumonia mentioned on the death certificate as a secondary cause of death (Gorina, Kelly, Lubitz, & Hines, 2008).

According to the CDC, more than 200,000 people are hospitalized in the United States each year as a result of complications from influenza (Centers for Disease Control and Prevention, 2008). The average hospital stay was approximately 5.3 days at a cost of $6,900 per stay (Milenkovic, Russo, & Elixhauser, 2006). Further, the death rate from influenza among the elderly ranges from 1.1 to 3.6 per 100,000, with risk increasing with age (Centers for Disease Control and Prevention, 2009). The death rate for influenza and pneumonia in people aged 65–74 years old is 3 times that of a person 55–64, and for a person over the age of 85, the death rate is 35 times higher (Centers for Disease Control and Prevention, 2009).

Among adults aged 65 years and older, approximately 67.4 percent were vaccinated during 2010, which is below the Health People 2020 target of 90 percent for this age group (Centers for Disease Control and Prevention; U.S. Department of Health and Human Services, 2000). For adults in nursing facilities, in 2004, the National Nursing Home Survey found that 62.9 percent of residents were vaccinated against influenza (National Nursing Home Survey, 2004). The most recent Nursing Home Compare data indicates that number has risen to 85 percent for short-stay residents and 92 percent for long-stay residents (Centers for Medicare & Medicaid Services, 2012).

This measure is intended to encourage nursing facilities to focus on this important aspect of clinical care by assessing residents on the status of their seasonal influenza vaccine immunization and to provide immunization as appropriate.

This is a very important measure of quality of care in nursing facilities. Morbidity and mortality related to influenza are often reported in conjunction with data regarding pneumonia, and together frequently lead to death in the elderly population. There is a demonstrated gap in performance in vaccination among adults aged 65 years or older.

There are two QMs for immunization-one for the long-stay and one for the short-stay populations. This is because the long-stay and short-stay populations are two distinct populations with inherent clinical differences that are aptly captured by these two length-of-stay categories. Furthermore, nursing facilities that serve a short-stay population have a limited time frame to assess and provide vaccinations compared to the long-stay population.

B. Differences between MDS 2.0 and MDS 3.0 QM Definitions

QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) reports the percent of short-stay nursing facility residents who are assessed and appropriately given the seasonal influenza vaccination during the influenza season as reported on the target MDS assessment: (1) received the influenza vaccine during the most recent influenza season, either in the facility (O0250A = 1) or outside the facility (O0250C = 2; (2) were offered and declined the influenza vaccine (O0250C = 4); or (3) were ineligible due to contraindications(s) (O0250C = 3). The numerator includes residents with a target assessment (OBRA MDS 3.0 assessment [A0310A = 01, 02, 03, 04, 05, 06], PPS assessment [A0310B = 01, 02, 03, 04, 05, 06], or discharge assessment [A0310.F = 10, 11]) during the influenza reporting period. When a vaccination is completed at the end of the influenza season, the next opportunity to report the vaccination may be after the season is over. Extending the seasonal influenza vaccination reporting period through 12 months allows for capturing those late season vaccinations reported after the season ends. Note that residents are not excluded from the measure calculation if there is a missing response on the item indicating whether the influenza vaccine was received (O0250A); rather, these residents are assumed to have not received the vaccine and counted in the denominator. Additionally, for residents who did not receive the vaccination (O0250A = 0), if the item indicating the reason it was not administered was left missing (O0250C), it is assumed that there was no valid reason for the resident not to receive the vaccine, and the resident is counted in the denominator of the measure.

The underlying MDS 3.0 items used to construct this measure did not change from the MDS 2.0 measure. Minor item changes to clarify the item included the addition of a "none of the above" category for the reason the vaccine was not given (O0250C = 9). In addition, the MDS 3.0 measure now includes individuals who refuse to be vaccinated or have medical contraindications to vaccination, potentially increasing the number of residents who might be counted in the aggregated QM numerator and denominator compared to the MDS 2.0 QM.

The short-stay population definition changed from MDS 2.0 to MDS 3.0. In the MDS 2.0 measure, residents were included in the post-acute care population if they had a 14-day PPS MDS in the target quarters; the measure was calculated based on just 14-day PPS assessments. In the MDS 3.0 measure, the analogous short-stay residents are defined as residents with 100 or fewer cumulative days in facility and whose assessments may be discharge assessments; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections summarize analyses using MDS 3.0 episode file for Quarter 4, 2011. Analyses address three general areas:

- Number of assessments that are included and excluded from the numerator of the QM based on sample restrictions.
- Whole sample prevalence of items that are related to the QM
- Findings regarding variability, reportability, reliability, and validity.

4.3.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

In Quarter 4, 2011, there were 1,068,388 assessments included in the denominator of QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) (approximately 71.7 percent of all assessments), and a total of 312,640 resident episodes (29.3 percent of the total number of short-stay resident episodes) were excluded from the denominator. ***Table 4.3-1*** shows the proportion of resident episodes excluded from the denominator because the residents were not in facility during the current or most recent influenza season, this measure's only exclusion criterion. A total of 312,640 (29.3 percent) assessments met this criterion.

Tables 4.3-1 and 4.3-2 illustrate the total target assessments not excluded from the denominator. Over 9 percent indicated that the resident had received the influenza vaccine during the most recent influenza season in the facility (O0250A = 1). Of those residents who did not receive the vaccine in the facility, 36 percent received the influenza vaccine outside the facility (O0250C = 2); 16 percent of residents were offered but declined the influenza vaccine (O0250C = 4); and a little, more than 1 percent of residents were ineligible to receive the vaccine due to contraindications(s) (O0250C = 3).

Table 4.3-1
Target assessments used in the calculation of QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	1,068,388	71.7%
Excluded—Exclusion 1	312,640	29.3
Total number of short-stay resident episodes	1,381,028	—

NOTES:

Exclusion 1 = resident not in facility during the current or most recent influenza season.

[a] Column values may not add up to total since a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident assessments in short-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

Other valid response categories to O0250C are missing [-]; not offered [5]; inability to obtain vaccine due to declared shortage [6]; or none of the above[9]. Frequencies associated with measure item categories are presented in ***Table 4.3-2***.

4.3.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with "perfect scores" (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes), which may indicate a "ceiling effect". ***Table 4.3-3*** shows the results of this analysis using the QM score for Quarter 4, 2011, and ***Tables 4.3-3a***, ***4.3-3b***, and ***4.3-3c*** show the results of all the submeasures for this QM. The mean score for QM #0680 is 79.8 percent with a standard deviation of 19.2 percent. This QM has an interquartile range of 23.3 percent. About 7.3 percent of facilities have scores of 100 percent ("perfect" scores).

Table 4.3-2
Frequency and percentage of responses on items used to define the numerator for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

QM #0680 Item	Entry	Frequency	Percentage	Cumulative percentage
O0250a	— = missing	56,475	4.1%	4.1%
O0250a	0 = No	1,199,343	86.8%	90.9%
O0250a	1 = Yes	125,210	9.1%	100.0%
O0250a	Total	1,381,028	—	—
O0250c	—	69,922	5.1%	5.1%
O0250c	1 = Resident not in facility during this year's flu season	312,640	22.6%	27.7%
O0250c	2 = Received outside of this facility	497,501	36.0%	63.7%
O0250c	3 = Not eligible – medical contraindication	16,080	1.2%	64.9%

(continued)

Table 4.3-2 (continued)
Frequency and percentage of responses on items used to define the numerator for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

QM #0680 Item	Entry	Frequency	Percentage	Cumulative percentage
O0250c	4 = Offered and declined	221,470	16.0%	80.9%
O0250c	5 = Not offered	78,290	5.7%	86.6%
O0250c	6 = Inability to obtain vaccine due to a declared shortage	1,513	0.1%	86.7%
O0250c	9 = None of the above	58,402	4.2%	90.9%
O0250c	^ = skipped	125,210	9.1%	100.0%
O0250c	Total	1,381,028	—	—

Analysis date: 4/17/2012

NOTE: O0250a = Resident received vaccine in-facility for the current influenza season); 0250c = If influenza vaccine not received, state reason.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_014_10.log)

Table 4.3-3
Score distribution for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Inter-quartile range
11,833	79.7%	19.2%	52.2%	70.7%	85.7%	94.0%	98.4%	7.3%	23.3%

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

Analysis date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_014_10.log)

Table 4.3-3a
Score distribution for submeasure QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
11,833	58.8%	21.5%	26.7%	44.6%	62.6%	75.1%	84.0%	30.5%

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

Analysis date: 7/3/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_014.log)

Table 4.3-3b
Score distribution for submeasure to QM #0680B Percent of Residents Who Were Offered and Declined the Seasonal Influenza Vaccine

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
11,833	19.3%	16.3%	3.8%	8.3%	15.1%	25.0%	40.9%	16.7%

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

Analysis date: 7/3/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_014.log)

Table 4.3-3c
Score distribution for submeasure QM #0680C Percent of Residents Who Did Not Receive, due to Medical Contraindication, the Seasonal Influenza Vaccine

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
11,833	1.6%	3.2%	0.0%	0.0%	0.0%	2.2%	4.2%	2.2%

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

Analysis date: 7/3/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_014.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 20 residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We examined the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.3-4*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 75.4 percent of facilities had sufficient sample size to meet the minimum requirements for public reporting QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay). This is comparable to the MDS 2.0 measure, which 75.7 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM calculation. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, slightly more than 75 percent of target assessments for the short-stay population for Quarter 4, 2011, were discharge assessments, and 8.8 percent of target assessments for the long-stay population in the same quarter were discharge assessments. This suggests that the inclusion of discharge assessments may affect the reportability of short-stay QMs to a greater degree than the long-stay QMs.

To evaluate the impact on reportability of including discharge assessments in the set of target assessments eligible for constructing QM #0680, we examined changes in the number of facilities able to report this QM when discharge assessments are not included in the set of target assessments. The results are presented in ***Table 4.3-4***. After excluding discharge assessments from the set of target assessments, reportability is slightly decreased for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) (75.4 percent with discharge assessments vs. 73.5 percent without discharge assessments).

Table 4.3-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (Percent) of resident episodes in the denominator	1,068,388 (77.4%)[a]	968,723 (70.1%)[a]
Number (Percent) of resident episodes in the numerator	860,261 (62.3%)[a]	778,381 (56.4%)[a]

(continued)

Table 4.3-4 (continued)
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (Percent) of facilities able to report this QM	11,833 (75.4%)[b]	11,534 (73.5%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the short-stay population (1,381,028).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of \MDS 3.0 episode file for Quarter 4, 2011(\quarter_4_5\db135_request\nh_014_10_nd.log and nh_014_10.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4, 2011.

Table 4.3-5 shows how facility rank changed from quarter to quarter in 2011. For each pair of quarters, large changes in ranking were rare. At least two-thirds of facilities stayed in approximately the same rank, and relatively few (no more than 7.0 percent) facilities had a rank change by more than three deciles.

Table 4.3-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 19.2 percent). Overall, the average change in scores for this measure was an increase of 2.6 percent. The vast majority of facilities had little or no change in their score from quarter to quarter (77.0 percent had changes of less than one standard deviation), whereas fewer than 2 percent had changes of more than three standard deviations.

Table 4.3-5
Distribution of change in facility rank from one quarter to the next for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	7,767 (67.5%)	6,744 (68.8%)	6,631 (67.9%)
Between 1 and 2 deciles	2,034 (17.7%)	1,597 (16.3%)	1,737 (17.8%)
Between 2 and 3 deciles	909 (7.9%)	753 (7.7%)	796 (8.2%)
More than 3 deciles	800 (7.0%)	709 (7.2%)	603 (6.2%)
Total	11,510 (100.0%)	9,803 (100.0%)	9,767 (100.0%)

NOTES;

Total number of facilities in each column reflects all the facilities that could report the measure in both quarters.

Analysis date: 6/15/2012

SOURCE: MDS 3.0, RTI analysis of MDS 3.0 episode files for Quarter 1 – Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.3-6
Change in facility score from one quarter to the next for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
9,767	2.6%	13.8%	0.4%	1.6%	8.0%	50.3%	26.7%	8.5%	3.4%	1.2%

NOTES:

Number of facilities reporting (9,767) reflects facilities that meet minimum requirements for public reporting this QM in both Quarter 3 and Quarter 4, 2011.

Analysis Date: 6/15/2012

SOURCE: RTI analysis of \ MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated). For example, one could argue that the "vaccination group" of measures—QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay), QM # 0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay), QM #0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay), and QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)—should reflect the quality of related clinical care processes. Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures. Among nursing homes that could report both related measures, we calculated the inter-QM correlation between the facility's percentile ranks on QM #0680, QM #0681, QM #0682, and QM #0683 for Quarter 4, 2011.[11] The correlation between the nursing home's percentile ranks on QM #0680 and QM #0681 was moderate ($r = 0.557$) and statistically significant ($p < .001$); the correlation between QM #0680 and QM #0682 was moderate ($r = 0.732$) and statistically significant ($p < .001$); and the correlation between QM #0680 and QM #0683 was also moderate ($r = 0.533$) and statistically significant ($p < .001$). The moderate correlation may indicate some facilities have systems in place for identifying patients in need of updating their vaccinations and providing them, whereas other facilities provide vaccinations on an as-needed or upon-request basis without a formal system in place.

Variation by State. For a quality measure to be valid, variation observed in the distribution of the QM should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by factors outside of the control of facilities, such as state-level payment policies or demographics, this can be a threat to the validity of the measure. To explore whether state characteristics might be a source of facility score variation for QM #0680 and thus a potential threat to the measure's validity we conducted ANOVA. We also examined the interquartile difference between the mean state-level scores across states. The proportion of variance in this measure explained by the state in which facilities are located is 4.0 percent and significant [$F(52, 11{,}780) = 9.5$, $p < .001$].[12] The difference between the mean state-level scores for states at the 25th percentile and the 75th percentile is 4.6 percentage points. Thus, the state in which a facility is located accounts for a significant but small proportion of the variance in the scores for this measure. While not a likely threat to validity, this issue merits future monitoring.

---

[11] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

[12] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

Missing Data. No short-stay resident was excluded from the QM calculation due to missing data, as values of zero are imputed for missing data on items used to calculate this quality measure.

In Quarter 4 of 2011, 29.3 percent of short-stay resident episodes were excluded from the calculation of the quality measure because they were not in the facility during the current or most recent influenza season (indicated by O0250C = [1]). Values of zero are imputed for missing data on items used to calculate this quality measure.

We examined the possible relationship between imputation rate and QM scores for this measure. The results are shown in ***Table 4.3-7***. There is a weak but significant correlation between imputation rate and QM scores for this measure ($r = 0.044$, $p < .001$). The result is counterintuitive, given that any missing data is counted as a "0" value and thus should suppress the total proportion of residents given the seasonal influenza vaccine. However, the positive relationship is very weak. Stratifying the mean facility-level imputation rates by QM score quartile shows that this relationship might not be linear: the means rise and fall across quartiles, and a relatively large proportion of facilities in the fourth quartile of QM scores have high imputation rates. However, imputation rate account for an extremely small proportion of the variance in QM scores (0.2 percent), and thus do not pose a substantial threat to validity.

Table 4.3-7
Distribution of facility-level imputation rate for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) by quartile of QM score

Quartile of QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,904	19.5%	16.7%	0.0%	3.3%	17.7%	31.5%	41.5%
26%–50%	3,903	21.9%	15.3%	0.0%	8.7%	22.0%	33.3%	41.6%
51%–75%	3,922	21.4%	14.9%	0.7%	8.4%	21.1%	32.8%	41.3%
76%–100%	3,884	20.7%	18.3%	0.0%	3.3%	18.2%	33.3%	45.0%

NOTES:

Total *n* (15,613) = number of facilities who have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in the overall impact of including the discharge assessment on QM rates and the completeness of the new discharge assessment items and its impact on the QM rate. The impact of the new discharge assessment on QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) was evaluated by examining each facility's QM score change before and after excluding discharge assessments

from the set of eligible target assessments. The results are presented in ***Table 4.3-8***. The mean nursing home–level QM score and distributions remain largely unchanged before and after excluding discharge assessments from the set of eligible target assessments. The mean score for this QM is 79.7 percent when discharge assessments are included in the set of target assessment and 80.0 percent when they are not (mean change = -0.3 percent). The use of the discharge assessment in the QM score calculations does not alter the distribution of scores among facilities.

Table 4.3-8
Distribution of facility-level score change after excluding discharge assessment for QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
11,522	-0.3%	4.6%	-4.8%	-2.2%	-0.1%	1.3%	3.9%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) after excluding discharge assessments.

Analysis date: 5/29/2012

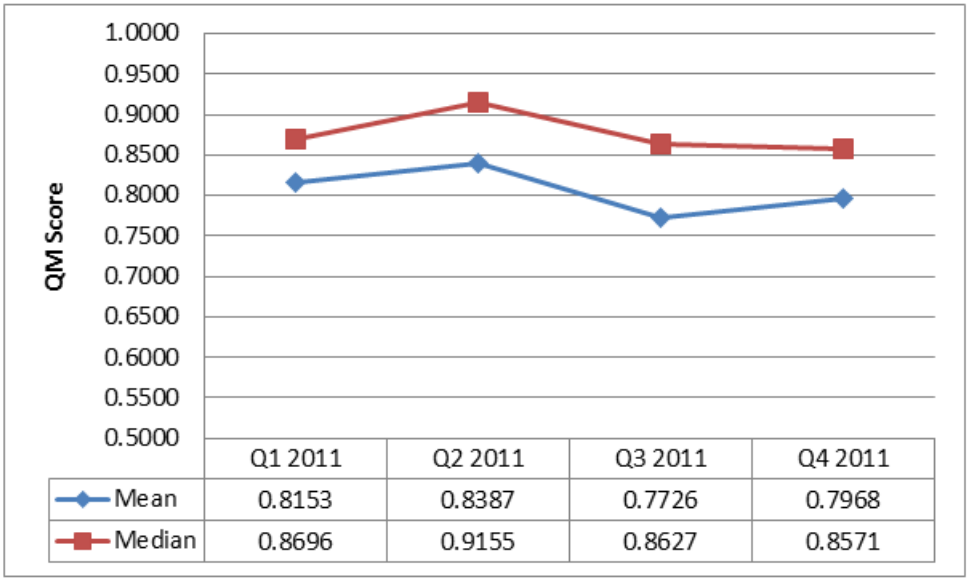
SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years, corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address whether seasonal variation might play a role in the score for QM #0680, we examined the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.3-1***.

The national-level means for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, were 81.5 percent, 83.9 percent, 77.3 percent, and 79.7 percent. The quarterly national medians were 87.0 percent, 91.6 percent, 86.3 percent, and 85.7 percent. Although the mean rose from Quarter 1, 2011, to Quarter 2, 2011; fell from Quarter 2, 2011, to Quarter 3, 2011; and rose again from Quarter 3, 2011, to Quarter 4, 2011, the median rose from Quarter 1, 2011, to Quarter 2, 2011, but fell for the rest of the year. With only 12 months of data available, it is too early to ascertain if the changes in national QM score over time reflect seasonal variation, and this will need to be confirmed when multiple years of data become available.

Figure 4.3-1
Seasonal (quarterly) variation in QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay)



SOURCE: analysis of MDS 3.0 episode files for Quarter 1–Quarter 4, 2011 \qm_quarter_1_2\complete\nh_014_10.log \qm_quarter_2_3\complete\nh_014_10.log \qm_quarter_3_4\complete\nh_014_10.log \qm_quarter_4_5\complete\nh_014_10.log

4.3.5 Risk Adjustment

QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) was endorsed by NQF without risk adjustment.

4.4 QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)

4.4.1 Summary of Findings

QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) reports the percentage of short-stay residents whose pneumococcal vaccination is up to date or who were offered and declined the vaccine or were ineligible due to contraindication(s). Because this QM does not have denominator exclusions, all short-stay residents are included in the denominator. Almost all facilities (99.7 percent) have 20 or more short-stay resident episodes included in the denominator and are able to report this QM. Among facilities able to report, the mean facility QM score is 81.0 percent. The QM score varies across facilities with a standard deviation of 21.0 percent and interquartile range of 23.0 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To examine the reliability of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. About half of facilities declined

in their score by less than one standard deviation and 31.2 percent of facilities increased their score by less than one standard deviation. A very low proportion of changes are greater than three standard deviations. A similar pattern was found for facility rank changes. About 80 percent of facilities remained within the same decile from quarter to quarter. Only about 2 percent of facilities shift more than 3 deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group (e.g., vaccination QMs) should be correlated with changes in other measures because they reflect similar care processes. Findings from this analysis showed that this QM is highly correlated with the long-stay pneumococcal vaccination QM ($r = 0.690$) and the short-stay influenza vaccination QM ($r = 0.732$). This QM is statistically significantly but moderately correlated with the long-stay influenza vaccination QM ($r = 0.342$; $p < .001$). We also analyzed the potential geographic (state) and seasonal (quarterly) variations in this QM. The state of the facility explains about 5 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

4.4.2 Background and Introduction to QM

This measure is based on data from MDS 3.0 assessments of short-stay nursing facility residents. The measure reports the percentage of all short-stay residents who were assessed and appropriately given the pneumococcal vaccination (PPV) as reported on the target MDS assessment (OBRA, PPS, or discharge) during the 12-month reporting period. This measure is harmonized with NQF's quality measure on pneumococcal immunizations (National Quality Forum, 2008). The MDS 3.0 definitions have been changed to conform to the NQF standard. The NQF used current guidelines from the Advisory Committee on Immunization Practices (ACIP) and others to guide decisions on all parameters for the harmonized measures (ACIP, 1997). The NQF standard specifications were harmonized to achieve a uniform approach to measurement across settings and populations, addressing who is included or excluded in the target denominator population, who is included in the numerator population, and time windows for measurement and vaccinations. Short-stay residents are those residents with 100 or fewer cumulative days in facility. The measure is restricted to the population that has short-term needs and does not include the population of residents with stays longer than 100 days. A separate quality measure is endorsed for the long-stay population.

A. QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) Background

This measure is intended to encourage nursing facilities to focus on this important aspect of clinical care by assessing residents on the status of their pneumococcal vaccine immunization and to provide immunization as appropriate. This is a very important measure of quality of care in the nursing facility. Morbidity and mortality related to pneumonia are often reported in conjunction with data regarding influenza, and together frequently lead to death in the elderly population. According to CDC, pneumonia and influenza were the seventh most common cause of death for persons aged 65 or older in the United States (Centers for Disease Control and Prevention, 2009). In 2009, pneumonia caused 50,774 deaths, 85 percent of which were in

people over the age of 65 (Centers for Disease Control and Prevention, 2009). In addition to being a primary cause of death, pneumonia and influenza are often complications of other illnesses. In 2004, there were approximately 123,000 deaths with influenza and pneumonia mentioned on the death certificate as a secondary cause of death (Gorina et al., 2008). The death rate from influenza and pneumonia in persons aged 65-74 years old is 3 times that of a person aged 55-64, and for a person over the age of 85, the death rate is 35 times higher (Centers for Disease Control and Prevention, 2009). According to CDC, pneumococcal disease kills more people in the United States each year than all other vaccine-preventable diseases combined (Centers for Disease Control and Prevention, 1997).

Older people and persons with chronic health conditions are at high risk for pneumococcal disease. However, there is a demonstrated gap in performance in vaccination among adults aged 65 years and older. In 2010, 68.6 percent of adults over the age of 65 reported having a pneumococcal vaccine (Centers for Disease Control and Prevention, 2010), less than the 90 percent goal set by Health People 2020 (U.S. Department of Health and Human Services, 2012). Vaccinations of nursing facility residents can prevent or lower the risk of residents becoming seriously ill. According to data from Nursing Home Compare, from 2007 to 2012, 91 percent of short-stay nursing home residents and 85 percent of long-stay residents have received the pneumococcal vaccine, up from 45.4 percent in 2004 (Centers for Medicare & Medicaid Services; National Nursing Home Survey, 2004).

Healthy People 2020 includes Objective IID-13, for institutionalized adults, of a 90 percent vaccination rate for pneumonia in 2020 (U.S. Department of Health and Human Services, 2012). Hospitalization rates for pneumonia-related stays for the elderly population have been increasing over the past 15 years, and among those 85 or older, at least 1 in 20 elderly persons were hospitalized each year because of pneumonia (Fry, Shay, Holman, Curns, & Anderson, 2005). In 2005, Medicare paid an average of $6,342 per hospital discharge for pneumonia-related short-stay hospitalizations; the average length of stay was 6.1 days. In 2007, There were 610,000 hospital discharges for pneumonia in patients over the age of 65, with an average length of stay of 5.0 days (Centers for Medicare & Medicaid Services, 2007).

B. Differences between MDS 2.0 to MDS 3.0 QM Definitions

QM #0682 reports the percentage of all short-stay residents who were assessed and appropriately given the seasonal pneumococcal vaccine during the influenza season. Residents are included in the numerator if they meet certain criteria on the target MDS 3.0 assessment (A0310A = 01,02,03,04,05,06; A0310B = 01,02,03,04,05,06; or A0310F = 10,11) with a reference date or discharge date during the 12-month reporting period, and (1) have an up-to-date PPV status (item O0300A = 1); or (2) were offered and declined the vaccine (item O0300B = 2); or (3) were ineligible due to medical contraindication(s) (i.e., anaphylactic hypersensitivity to components of the vaccine, bone marrow transplant within the past 12 months, or receiving a course of chemotherapy within the past 2 weeks) (item O0300B = 1).

There are no resident-level exclusions from the denominator. The denominator includes all short-stay residents who meet the following criteria: (1) the target MDS 3.0 assessment is an OBRA assessment (item A0310A = 01, 02, 03, 04, 05, 06) or a PPS assessment (A0310B = 01, 02, 03, 04, 05, 06) or a discharge assessment (item A0310F = 10,11) with discharge date (item

A2000) during the 12-month target period. Note that residents are not excluded from the measure if there is a missing response on the item indicating whether their PPV status is up to date (O0300A); rather, it is assumed that these residents do not have an up-to-date PPV vaccination and they are counted in the denominator. Additionally, for residents who do not have an up-to-date PPV vaccination (O0300A = 0), if the item indicating the reason it was not administered was left missing (O0300B), it is assumed that there was no valid reason for the resident not to receive the vaccine and the resident is counted in the denominator of the measure.

This measure is based on the NQF's National Voluntary Standards for Influenza and Pneumococcal Immunizations, which include resident refusal and ineligibility in the numerator and denominator. This is a change from the MDS 2.0 nursing home quality measure. The underlying MDS items used to construct this measure did not change from the MDS 2.0 to MDS 3.0. The changes made to the MDS 3.0 regarding the vaccine items were relatively minor wording changes meant to clarify the item. The reasons for the pneumococcal vaccine not being received are identical to the MDS 2.0 item. Finally, as with the MDS 2.0 pneumococcal vaccination measures, the MDS 3.0 calculates separate measures for the short and long-stay populations. There are inherent differences in nursing facility's being responsible for assessing and/or providing vaccines for these distinct populations. For the short-stay population, nursing facilities have less time to assess and/or provide the vaccine than for the long-stay population. As a result, nursing facilities' vaccination rates for post-acute care populations should not be compared to rates for long-term care populations. Separating them recognizes these differences in vaccination rates.

Note that the main difference between the MDS 2.0 and MDS 3.0 measures is that the overall definitions changed. In the MDS 2.0 measure, residents were included in the post-acute care population if they had a 14-day PPS MDS in the target quarters; the measure was calculated based on just 14-day PPS assessments. In the MDS 3.0 measure, the analogous population is the short-stay residents, defined as residents with 100 or fewer cumulative days in facility, and whose assessments may be discharge, 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments, or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections present analyses using MDS 3.0 episode file for Quarter 4, 2011(unless otherwise specified). Analyses address three general areas:

- Number of assessments that are included and excluded from the numerator of the QM based on sample restrictions
- Whole sample prevalence of items that are related to the QM
- Findings regarding variability, reportability, reliability, and validity

4.4.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 1,381,028 assessments (100 percent) included in the denominator of QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) and no resident episodes were excluded from the denominator of QM #0682, because this measure has no exclusion criteria.

Table 4.4-1
Target assessments used in the calculation of QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage
Included	1,381,028	100.0%
Total number of short-stay resident episodes	1,381,028	—

NOTES:

There are no exclusions for this measure.

[a] Column values may not add up to total since a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident assessments in short-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

Table 4.4-2 describes the responses to the key MDS 3.0 item used to calculate this QM. Item O0300a asks whether the resident's pneumococcal vaccination is up to date. Nearly 59 percent of residents indicate that the vaccination is up to date (O0300a = 1) and would be counted in the numerator based on this information. This is lower than among long-stay residents (QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine [Long Stay]). About 23 percent of residents were offered the vaccine and declined (O300b = 2) which would also count them in the numerator. This is higher than among long-stay residents.

Table 4.4-2
Responses to O0300a (PPV status is up to date) and O0300b (reason for not receiving pneumonia vaccine) on target assessments

Item	Entry	Frequency	Percentage	Cumulative %
O0300a	- = blank	90,267	6.5%	6.5%
O0300a	0 = not up to date	476,610	34.5	41.1
O0300a	1 = up to date	814,151	59.0	100.0
O0300a	Total	1,381,028	—	—

(continued)

Table 4.4-2 (continued)
Responses to O0300a (PPV status is up to date) and O0300b (reason for not receiving pneumonia vaccine) on target assessments

Item	Entry	Frequency	Percentage	Cumulative %
O0300b	— = missing	124,045	9.0	9.0
O0300b	1= not eligible—medical contraindication	23,477	1.7	10.7
O0300b	2 = offered and declined	313,323	22.7	33.4
O0300b	3 = not offered	106,032	7.7	41.1
O0300b	^ = skipped because vaccine is up to date	814,151	59.0	100.0
O0300b	Total	1,381,028	—	—

NOTE: O0300a = Pneumococcal vaccination up to date; O0300b = If Pneumococcal vaccine not received, state reason.

Analysis date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_016_10.log)

4.4.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with "perfect scores" (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes) which may indicate a "ceiling effect". ***Table 4.4-3*** shows the results of this analysis using the QM score for Quarter 4, 2011, and ***Tables 4.4-3a***, ***4.4-3b***, and ***4.4-3c*** show the results for the submeasures. The mean score for QM #0682 is 81.0 percent with a standard deviation of 21.0 percent. This QM has an interquartile range of 23.0 percent. About 9.9 percent of facilities have scores of 100 percent ("perfect" scores).

Table 4.4-3
QM score distribution for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Inter-quartile range
12,762	81.0%	21.0%	50.0%	73.2%	88.6%	96.2%	99.7%	9.9%	23.0%

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011(.\qm_quarter_4_5\complete\nh_016_10.log)

Table 4.4-3a
QM score distribution for submeasure QM #0682A Percent of Residents Who Received the Pneumococcal Vaccine (Short Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
12,762	59.1%	24.3%	21.4%	42.6%	64.1%	78.2%	87.2%	35.6%

Analysis date:7/3/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

SOURCE: RTI Analysis of MDS 3.0 episode file for Quarter 4, 2011 (.\quarter_4_5\db178_request\db178_request_016.log)

Table 4.4-3b
QM score distribution for submeasure QM #0682B Percent of Residents Who Were Offered and Declined the Pneumococcal Vaccine (Short Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
12,762	20.3%	18.7%	2.9%	7.4%	14.8%	27.0%	46.0%	19.7%

Analysis date:7/3/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

SOURCE: RTI Analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_016.log)

Table 4.4-3c
QM score distribution for submeasure QM #0682C Percent of Residents Who Did Not Receive, due to Medical Contraindication, the Pneumococcal Vaccine (Short Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
12,762	1.6%	4.2%	0.0%	0.0%	0.0%	1.9%	4.2%	1.9%

Analysis date:7/3/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

SOURCE: RTI Analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_016.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 20 residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examine the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.4-4*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 99.7 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) using MDS 3.0. This is an improvement over the MDS 2.0 measure, where 76.0 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM calculation. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay population for Quarter 4, 2011, were discharge assessments, and 8.8 percent of target assessments for long-stay residents in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may affect the reportability of short-stay QMs to a greater degree than the long-stay QMs.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0682, we examined changes in the number of facilities able to report this QM when discharge assessments are not included in the set of target assessments. The results are presented in ***Table 4.4-4***. After excluding discharge assessments from the set of target assessments, reportability slightly decreased for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) (81.0 percent with discharge assessments vs. 80.2 percent without discharge assessments).

Table 4.4-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (Percent) of resident episodes in the denominator	1,381,028 (100.0%)[a]	1,270,753 (92.0%)[a]
Number (Percent) of resident episodes in the numerator	1,150,951 (83.3%)[a]	1,060,255 (76.8%)[a]
Number (Percent) of facilities able to report this QM	12,762 (81.0%)[b]	12,578 (80.2%)[b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the short-stay population (1,381,028).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay residents (15,686).

Analysis date: 5/29/2012

SOURCE: MDS 3.0, RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\nh_016_10_nd.log and nh_016_10.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.4-5 shows how facility ranks on this quality measure changed from quarter to quarter in 2011. For each pair of quarters, large changes in ranking were rare: at least two-thirds of facilities stayed in approximately the same rank, and few (no more than 2.5 percent) facilities saw their rank change by more than three deciles.

Table 4.4-6 presents the results from an analysis of the change in facility scores for this quality measure between quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 21.0 percent). Overall, the average change in scores for this measure was an increase of 0.9 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (78.3 percent saw changes of less than one standard deviation), while slightly less than 2 percent saw changes of more than three standard deviations.

Table 4.4-5
Distribution of change in facility rank from one quarter to the next for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	9,908 (79.8%)	10,125 (81.2%)	10,221 (82.4%)
Between 1 and 2 deciles	1,607 (12.9%)	1,611 (12.9%)	1,539 (12.4%)
Between 2 and 3 deciles	591 (4.8%)	464 (3.7%)	422 (3.4%)
More than 3 deciles	310 (2.5%)	269 (2.2%)	223 (1.8%)
Total	12,416	12,469	12,405

Analysis date: 6/15/2012

NOTES:

Total *n*'s reflect facilities that meet minimum requirements for public reporting this QM in both Quarters.

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1–Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.4-6
Changes in facility scores from one quarter to the next for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
12,405	0.9%	8.2%	0.8%	1.9%	7.1%	47.1%	31.2%	7.6%	2.4%	1.1%

Analysis Date: 6/15/2012

NOTES:

Number of facilities reporting (12,405) reflects facilities that meet minimum requirements for public reporting this QM in both Quarter 3 and Quarter 4, 2011.

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated). For example, one could argue that the "vaccination group" of measures—QM #0680 Percent of Nursing Home Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay), QM # 0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay), QM #0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay), and QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)—should reflect the quality of related clinical care processes. Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures, where applicable. Among nursing homes that could report both related measures, we calculated the inter-QM correlations between the facility's percentile rank on QM #0682, QM #0680, QM #0681, and QM #0683 for Quarter 4, 2011.[13] We found that the correlation between the nursing home's percentile ranks on QM #0682 and QM #0680 was moderate ($r = 0.732$) and statistically significant ($p < .001$); the correlation between QM #0682 and QM #0681 was moderate ($r = 0.342$) and statistically significant ($p < .001$); and the correlation between QM #0682 and QM #0683 was also moderate ($r = 0.690$) and statistically significant ($p < .001$). The moderate correlation possibly indicates that some facilities have systems in place for identifying patients in need of updating their vaccinations and providing them, whereas other facilities provide vaccinations on an as-needed or upon-request basis without a formal system in place.

Variation by State. For a quality measure to be valid, variation observed in the distribution of the QM should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as state-level payment policies or demographics, this can be a threat to the validity of the measure. To explore whether state characteristics might be a source of facility score variation for QM #0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay) and thus a potential threat to the measure's validity, we conducted ANOVA. We also examined the interquartile difference between the mean state-level scores across states. The proportion of variance in this measure explained by the state in which facilities are located is 4.8 percent and significant [F (52, 10,980) = 16.5, $p < .001$].[14] The difference between the mean state-level scores for states at the 25th percentile and the 75th percentile is 6.8 percentage points. Thus, while the majority of the variance in QM #0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay)

---

[13] RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

[14] SOURCE: RTI analysis of MDS 3.0 data for Quarter 4, 2011(\quarter_4_5\db160_request)

scores is due to factors other than geography, there is a small and significant proportion of the variance that is explained by the respective states in which nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. No short-stay resident was excluded from the QM calculation due to missing data.

By definition, this measure includes all short-stay residents in its denominator and imputes values of [0] for missing data. Thus, all 1,381,028 short-stay residents were included in the denominator for this measure in quarter 4 of 2011.

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in the overall impact of including the discharge assessment on QM rates and the completeness of the new discharge assessment items and its impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) by comparing distributions of nursing home level QM scores when calculating the QM with and without discharge assessments in the set of target assessments. The results are presented in ***Table 4.4-7.*** The mean nursing home level QM score and distributions change slightly before and after excluding discharge assessments from the set of eligible target assessments. The mean score for this QM changed by less than one tenth of one percent when discharge assessments were included. The use of the discharge assessment in the QM score calculations does not alter the distribution of scores among facilities.

Table 4.4-7
Distribution of facility-level score change after excluding discharge assessment for QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
12,578	0.0%	3.7%	-3.4%	-1.2%	0.0%	1.0%	3.3%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

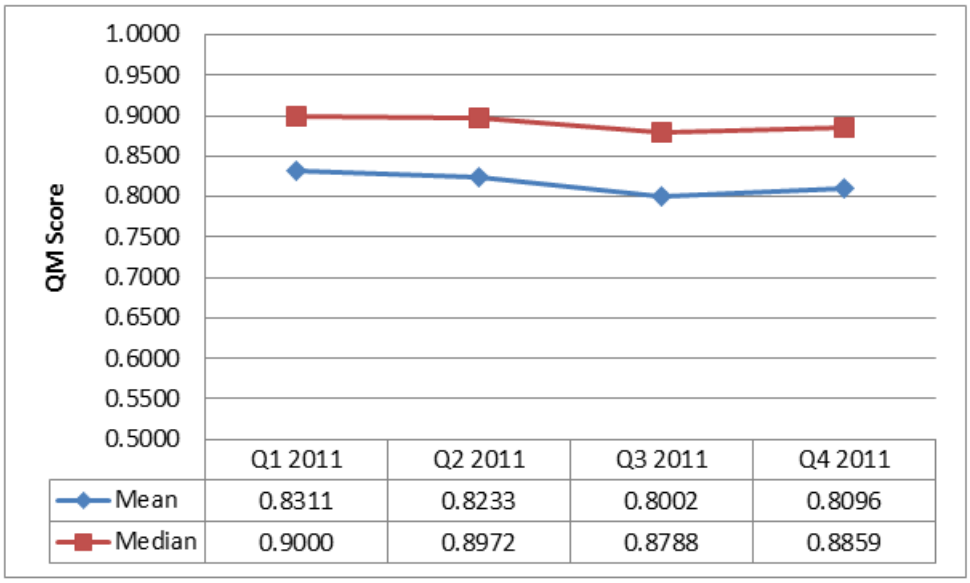
Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To explore whether

seasonal variation might play a role in the score for QM #0682, we examined the national mean and median for this QM score between Quarter 1, 2011 and Quarter 4, 2011. The results are presented in ***Figure 4.4-1.***

The national-level means for this measure for each quarter between Quarter 1, 2011–Quarter 4, 2011 were 83.1 percent, 82.3 percent, 80.0 percent, and 81.0 percent. The quarterly national medians were 90.0 percent, 89.7 percent, 87.9 percent, and 88.6 percent. Both mean and median remained relatively constant, with a small dip in Quarter 3, 2011. With only 12 months of data available, it is too early to ascertain if the changes in national QM score over time reflect seasonal variation; this will need to be confirmed when multiple years of data become available.

Figure 4.4-1
Seasonal (quarterly) variation in QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay)



SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1–Quarter 4, 2011 \qm_quarter_1_2\complete\nh_016_10.log; \qm_quarter_2_3\complete\nh_016_10.log ;\qm_quarter_3_4\complete\nh_016_10.log; \qm_quarter_4_5\complete\nh_016_10.log

4.4.5 Risk Adjustment

QM #0682 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Short Stay) was endorsed by NQF without denominator exclusion and model-based risk adjustment.

4.5 QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)

4.5.1 Summary of Findings

QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) reports the percentage of long-stay residents who have one or more falls that

resulted in major injury during the reporting period. We found that all long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 88.4 percent of the facilities have 30 or more long-stay resident episodes included in the denominator and are able to report this QM. Among facilities able to report, the mean facility QM score is 3.4 percent. The QM score varies across facility with a standard deviation of 2.7 percent and interquartile range of 3.3 percent. The relatively small variability of this QM is largely due to the low incidence rate.

To determine the reliability of this QM, we examined changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation; a very low proportion of changes were greater than three standard deviations. A similar pattern was found for facility rank changes. About 60 percent of facilities remained within the same decile from quarter to quarter. About 90 percent of facilities have rank changes within three deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a care process group. The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures because they reflect similar care processes. We selected QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the two QMs (correlation $r = 0.111$). Second, we examined the frequency and distribution of missing data. Missing rate is very low (nearly zero) for the items used to construct this QM and therefore should not pose a threat to validity of the QM. Last but not least, we also analyzed the potential geographic (State) variations in this QM. The State in which a facility is located explains about 7.8 percent of the variation in this QM. RTI will examine this issue further when more data become available.

As reliability and validity tests support acceptable reliability and validity of this QM, QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) received full NQF endorsement on August 1, 2012.

4.5.2 Background and Introduction to QM

This measure is based on data from all MDS 3.0 assessments of long-stay nursing facility residents, which may be Omnibus Budget Reconciliation Act of 1987 (OBRA) admission, annual, quarterly, significant change, or significant correction; PPS 5-day, 14-day, 30-day, 60-day, 90-day, or readmission/return; or discharge assessment. It reports the percentage of residents who experienced one or more falls with major injury (i.e., bone fractures, joint dislocations, closed head injuries with altered consciousness, and subdural hematoma) in the last year (12-month period). The measure is based on MDS 3.0 item J1900C, which indicates whether any falls that occurred were associated with major injury.

A. Background for This QM

Research findings indicate that approximately 75 percent of nursing facility residents fall at least once a year—twice the rate of their counterparts in the community (Rubenstein,

Josephson, & Robbins, 1994). Falls are the leading cause of unintentional injury in adults over the age of 65, resulting in 15,800 deaths, 1.8 million visits to U.S. emergency departments, and more than 300,000 hospitalizations (Centers for Disease Control and Prevention, 2005; Stevens, Thomas, Teh, & Greenspan, 2009). This is equivalent to one emergency room visit every 17 seconds, and one death every 30 minutes (Centers for Disease Control and Prevention, 2005). Further, it is estimated that 10 percent to 25 percent of nursing facility resident falls result in fractures and/or hospitalization (Vu, Weintraub, & Rubenstein, 2004). Saliba and Buchanan tested the proposed MDS 3.0 items, including those assessing the prevalence of any falls and falls with major injuries. Their study included 4,586 residents from 71 community nursing facilities and 19 Veteran's Administration nursing facilities in eight different States and found rates of falls and falls with injury similar to those reported in the literature. During their 6-month data collection period, they found that approximately 24 percent of patients reported at least one fall since the prior assessment. Among the 24 percent who experienced a fall, 9 percent of them had at least one fall with major injury and an additional 30 percent of them had at least one fall with minor injury (Saliba & Buchanan, 2008). The effort to monitor the prevalence of injurious falls at the facility level is very important for protecting the health of residents in the nursing facility. Research has shown that falls resulting in serious injury, such as hip fracture, are a leading cause of death and disability in this population, and are associated with excess mortality, functional limitations, loss of independence, and decreased quality of life (Rubenstein et al., 1994; Stevens et al., 2009). Moreover, studies show that such falls can leave up to 50 percent to 65 percent of residents with fears of falling (Magaziner et al., 1997) that affect both their functional abilities and social activities (Yardley & Smith, 2002).

Falls in the elderly represent a significant cost burden to the entire health care system. The average cost of a hospitalization for adults over the age of 65 is $18,667, resulting in an annual cost of approximately $5.8 billion; the average cost of treatment in an Emergency Department is $1,093, with an annual cost of approximately $1.1 billion; and the average cost of a fall resulting in a fatality is $22, 114, resulting in an annual cost of approximately $350 million (all figures in 2005 dollars; Centers for Disease Control and Prevention, 2005). Other estimates place the total cost of falls at $23.6 billion dollars (Centers for Disease Control and Prevention, 2005). Sorensen et al. (2006) document the costs associated with falls of varying severity in acute care hospitals, as well as in post-acute and long-term care settings. Their work suggests that in acute care hospitals, the costs incurred for falls with major injury range from $979 for a typical case with a simple fracture to $14,716 for a typical case with multiple injuries (Sorensen et al., 2006). In addition, Sorensen et al. also compared the cost of post-falls care in long-term care settings with the cost of patient care prior to hospitalization. They found that patients who experienced a fall with major injury had an additional post-fall Resource Utilization Group (RUG) III per diem cost ranging from $1,999 per year for a typical case with a simple fracture to $15,992 per year for a case with multiple injuries (Sorensen et al., 2006).

B. Summary of Differences between MDS 2.0 to MDS 3.0 QM Definitions

The measure of falls with major injury in the MDS 3.0 is a new measure, based on new items from the updated MDS 3.0 item set— J1900c, which measures whether the resident experienced one or more falls with major injury while in the facility. The numerator is based on the number of long-stay nursing facility residents who experienced one or more falls that resulted in major injury (J1900C = 1 or 2) on any assessment in the last 12 months. In the MDS 3.0,

major injury is defined as bone fractures, joint dislocations, closed head injuries with altered consciousness, or subdural hematoma. The denominator is the total number of long-stay nursing home residents, excluding those where *all* their MDS 3.0 assessments in the look-back period do not contain information in Section J (i.e., the occurrence of falls was not assessed (J1800 is missing) or the assessment indicates that a fall occurred (J1800 = [1]) AND the number of falls with major injury was not assessed (J1900C = [-]).

The similar item in the MDS 2.0 was item J4, which captured whether the resident experienced any falls within the past 30 days. The new item is an improvement in that it captures only falls that occurred during the stay within the facility, and differentiates between falls resulting in no injury, falls resulting in minor injury, and falls resulting in major injury. Falls that occurred prior to admission to the facility are not counted in this measure, an improvement over the previous items in MDS 2.0. This new quality measure using MDS 3.0 identifies those facilities with residents experiencing falls with major injury, defined as bone fractures, joint dislocations, and closed head injury with altered consciousness, and subdural hematoma. This new measure is an improvement over the MDS 2.0 measure in that it allows for a more specific assessment of facility safety.

In addition, the overall short-stay and long-stay definitions changed from MDS 2.0 to MDS 3.0. For the MDS 2.0, residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments and typically capture residents with a 30-day average length of stay. For the MDS 3.0, the analogous sample is for long-stay residents, defined as residents with more than 100 cumulative days in the facility, and assessments may be discharge; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections summarize the basic analyses performed for this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- number of assessments that are included and excluded from the numerator of the QM based on sample restrictions;
- whole sample prevalence of items that are related to the QM; and
- findings regarding variability, reportability, reliability, and validity.

4.5.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 1,203,663 (100 percent) episodes included in the numerator of QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay), and a total of 63 resident episodes (0.0 percent of the total number of long-stay resident episodes) were excluded from the denominator. ***Table 4.5-1*** shows the proportion of resident episodes excluded from the denominator for this measure's exclusion criterion. For this measure, the only criterion is missing data: residents are excluded from the denominator if

the occurrence of falls was not assessed or if the assessment indicates that the resident did fall but the number of falls with major injury was not assessed. A total of 63 (0.0 percent) assessments met this criterion. Overall, 1,203,663 assessments were included in the denominator (approximately 100 percent of all assessments).

Table 4.5-1
Target assessments used in the calculation of QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	1,203,663	100.0%
Excluded—Exclusion 1	63	0.0%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Missing Data

[a] Column values may not add up to total because a resident episode can meet more than one exclusion criterion. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

Table 4.5-2 below shows the percentage of residents with at least one assessment in the look-back period indicating a fall with major injury, thus triggering the QM. The vast majority of the resident episodes did not have an assessment in the look-back period that triggers the QM, that is, indicating a fall resulting in major injury. Only 39,158 resident episodes (3.3 percent) had at least one assessment that triggered the numerator.

Table 4.5-2
Percentage of resident episodes in which at least one assessment in the look-back period indicated a fall resulting in major injury

QM numerator	Number	Percent	Cumulative percent
No assessment in look-back period triggered the QM	1,164,505	96.8%	96.8%
At least one assessment in look-back period triggered the QM	39,158	3.3%	100.0%

NOTE: An assessment triggered the QM if the assessment indicated a fall resulting in major injury.

Analysis date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_008_10.log)

4.5.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with "perfect scores" (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes), which may indicate a "ceiling effect." ***Table 4.5-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) is 3.4 percent with a standard deviation of 2.7 percent. This QM has an interquartile range of 3.3 percent. About 14.4 percent of facilities have scores of 0 percent (perfect scores).

Table 4.5-3
QM score distribution for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)**

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with perfect Scores	Interquartile range
13,868	3.4%	2.7%	0%	1.5%	2.9%	4.8%	6.8%	14.4%	3.3%

NOTE: **Measure is calculated over a 12-month period.

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting of this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_008_10.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 long-stay residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examined the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.5-4*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 88.4 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) using the MDS 3.0. This measure was not calculated using the MDS 2.0 measure, so there is no way to compare results from facilities in Quarter 1, 2006.

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and that 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may improve the reportability of long-stay QMs to some extent, though the impact is likely less than for the short-stay QMs. To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay), we examined changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.5-4.*** After excluding discharge assessments from the target assessment sample, reportability is largely unchanged for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) (88.42 percent with discharge assessments vs. 88.36 percent without discharge assessments). Thus, the inclusion of the discharge assessment had virtually no effect on the number of facilities able to report this measure.

Table 4.5-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator	1,203,633 (100.0%) [a]	1,202,266 (99.9%) [a]
Number (%) of resident episodes in the numerator	39,158 (3.3%) [a]	27,067 (2.2%) [a]
Number (%) of facilities able to report this QM	13,868 (88.4%) [b]	13,860 (88.4%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay population (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.5-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. For each pair of quarters, the majority of facilities stayed in approximately the same rank position. Relative to other measures, a large number of facilities (more than 10 percent) saw changes of more than three deciles in each pair of quarters. However, this is a low-prevalence measure, in which changes of a few percentage points may lead to large changes in rankings. When taking this scale of scores into account (as we do in the next analysis), it is easy to see that they are not changing very much from quarter to quarter.

Table 4.5-5
Distribution of change in facility rank from one quarter to the next, QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	8,527 (62.6%)	8,417 (61.3%)	7,995 (58.2%)
Between 1 and 2 deciles	2,324 (17.1%)	2,319 (16.9%)	2,591 (18.9%)
Between 2 and 3 deciles	1,240 (9.1%)	1,537 (11.2%)	1,556 (11.3%)
More than 3 deciles	1,521 (11.2%)	1,451 (10.6%)	1,587 (11.6%)
Total	13,612	13,724	13,729

NOTES:

Total *n*'s reflect facilities that meet minimum requirements for public reporting this QM in both quarters.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 1 through Quarter 4, 2011 (\\quarter_4_5\db155_request\db155_request.log)

Table 4.5-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 2.7 percent). Overall, the average change in scores for this measure was a decline of 0.1 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (72.8 percent saw changes of less than one standard deviation), whereas just over 1 percent saw changes of more than three standard deviations.

Table 4.5-6
Change in facility score from one quarter to the next, QM #0674 Percent of Residents Experiencing One of More Falls With Major Injury (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
13,729	-0.1%	1.9%	0.6%	2.0%	11.0%	28.7%	44.1%	10.8%	2.2%	0.7%

NOTES:

Number of facilities reporting (13,729) reflects facilities that meet minimum requirements for public reporting this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures where applicable. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) and the facility's percentile rank on QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) for Quarter 4, 2011 These two measures are related in that they have the physical abilities of residents as an outcome characteristic. We found that the correlation between the nursing home's percentile rank on QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) and percentile rank on QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) was small (correlation coefficient = 0.111) but statistically significant ($p < 0.001$).[15]

Variation by State. For a quality measure to be valid, variation observed in the distribution of the QM should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether State characteristics might be a source of facility score variation for QM # 0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay), we conducted a one-way analysis of variance (ANOVA). We also examined the interquartile difference between the mean State-level scores across States. The proportion of variance in this measure explained by the State in which facilities are located is 7.8 percent and significant [$F(51, 13{,}816) = 22.9$, $p < .001$].[16] The difference between the mean State-level scores for States at the 25th percentile and the 75th percentile is 1.1 percentage points. Thus, although this measure varies little between States, a reliable but small proportion of the variance can be explained by geography, an issue that should be monitored in the future.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns of missing data indicate that certain types of residents tend to have assessments with missing data in ways that affect the calculation of a quality measure, then that measure may not be capturing processes and outcomes of care for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

---

[15] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db173_request\db173_request.log)

[16] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

In Quarter 4 of 2011, 63 residents (less than one-tenth of 1 percent of the long-stay population) were excluded from the calculation of this measure, all due to missing data on the items related to falls and subsequent injuries (J1800 and J1900C).

Table 4.5-7 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. Having missing data is exceedingly rare on items used to construct this measure: at least 90 percent of facilities have no missing data. Thus, there is clearly no threat to validity stemming from incomplete data for items used to calculate this measure.

We further examined the possible relationship between missing data and QM scores. ***Table 4.5-8*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure; they are essentially identical across quartiles. Further, there is no correlation between missing data and QM scores (r = -0.009, *n.s.*). Because there were so few missing data for this measure, there was no discernible relationship between the missing rate and the QM score.

Table 4.5-7
Distribution of facility-level missing rate for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,399	0.0%	0.9%	0%	0%	0%	0%	0%

NOTES:

n = number of facilities that have data for the numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request.log)

Table 4.5-8
Distribution of facility-level missing rate for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) by Quartile of QM score

Quartile of QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
1–25	3,868	0.0%	0.8%	0%	0%	0%	0%	0%
26%–50%	3,898	0.0%	0.1%	0%	0%	0%	0%	0%
51%–75%	3,795	0.0%	0.2%	0%	0%	0%	0%	0%
76%–100%	3,837	0.0%	0.1%	0%	0%	0%	0%	0%

NOTES:

Total *n* = number of facilities that have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in one, the overall impact of including the discharge assessment on QM rates and two, the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.5-9***. The mean nursing home–level QM score and distributions changed slightly before and after excluding discharge assessments from the set of eligible target assessments. The average facility score decreased by 1.0 percentage point when discharge assessments were excluded, suggesting a third of the falls were in the discharge assessment (these falls may be the reason for discharging the residents to a hospital). Thus, including the discharge assessment captures a number of people who would otherwise be missed. The impact of discharge assessments on the QM score varied to some extent, but most facilities did not have score change (median change = 0 percent).

Table 4.5-9
Distribution of facility-level score change after excluding discharge assessment for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
13,860	-1.0%	1.4%	-2.9%	-1.8%	0.0%	0.0%	0.0%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

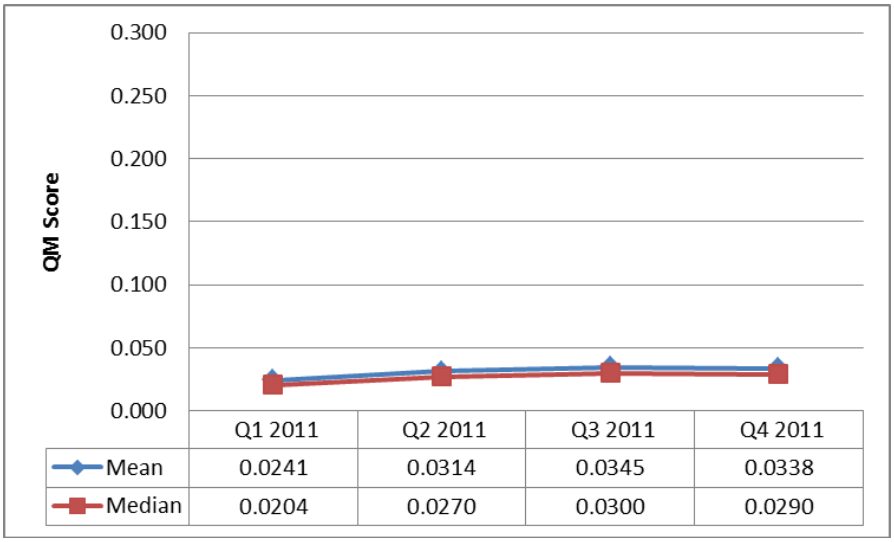
SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request_v1\db135_request_v1.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years, corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address whether seasonal variation might play a role in the score for QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay), we examined the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.5-1***.

The national-level means for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, were 2.4 percent, 3.1 percent, 3.5 percent, and 3.4 percent. The quarterly national medians were 2.0 percent, 2.7 percent, 3.0 percent, and 2.9 percent. Both the mean and median rose from Quarter 1, 2011, through to Quarter 3, 2011, and fell from Quarter 3, 2011, to Quarter 4, 2011.

Figure 4.5-1
Seasonal (quarterly) variation in QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay)



4.5.5 Risk Adjustment

QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) was endorsed by NQF without denominator exclusion and model-based risk adjustment. The review of the literature and clinical expert feedback recommended that facilities be held responsible for monitoring and preventing falls for high-risk residents.

4.6 QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

4.6.1 Summary of Findings

QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) reports the percentage of long-stay residents who self-report daily pain and at least one episode of moderate/severe pain or very severe/horrible pain of any frequency. This QM is risk adjusted using indirect standardization, adjusting for one covariate based on the resident's prior assessment in the episode. The covariate has a value of 1 if the resident shows independence or modified independence in daily decision making (MDS 3.0 item C1000 = [0,1]) or has a Brief Interview for Mental Status (BIMS) score from 13 to 15.

We found that 71.8 percent of the long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, more than 70 percent of the facilities had 30 or more long-stay resident episodes included in the denominator and were able to report this QM. Among facilities able to report, the mean facility risk-adjusted QM score was 11.5 percent. The QM score varied across facilities with a standard deviation of 8.5 percent and interquartile range of 11.5 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To determine the reliability of this QM, we examined the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation—a very low proportion of changes were greater than three standard deviations. About 60 percent of facilities remained within the same decile from quarter to quarter. About 7 percent of facilities had rank changes of more than three deciles from quarter to quarter in 2011.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a care process group. The hypothesis was that changes in one QM within a measure group (e.g., short-stay and long-stay pain QMs) should be correlated with changes in other measures because they reflect similar care processes. The findings from this analysis showed significant and moderate correlations between the short-stay and long-stay pain QMs (correlation $r = 0.560$). Second, we examined the frequency and distribution of missing data. We found that missing rate varies by some resident characteristics; for example, older residents (85 years old or older) or residents with cognitive impairment were more likely to have missing data. The correlation analysis showed that the correlation between the missing rate and the QM score was very weak, indicating that missing data should not pose a threat to validity of the QM. Finally, we also analyzed the potential geographic (State) and seasonal (quarter) variations in this QM. The State in which the facility is located explains about 9.5 percent of the variation in this QM. The QM score remained largely stable from Quarter 1 to 4 in 2011.

This QM uses model-based risk adjustment adjusting for one covariate, which is independence or modified independence in daily decision making on the prior assessment. We examined the impact of risk adjustment on the QM scores and facility score changes. We also compared coefficients for the covariates across quarter. The results support the validity of the risk-adjustment model. In addition, we explored different model specifications (single-level logistic regression models vs. hierarchical models) for the risk-adjustment model. The results show that changing model specifications has a big impact on facility rank based on the risk-adjusted QM score.

As reliability and validity tests support acceptable reliability and validity of this QM, #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay), received full NQF endorsement on August 1, 2012.

4.6.2 Background and Introduction to QM

QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) reports the percentage of long-stay residents who report (1) almost constant or frequent pain and at least one episode of moderate to severe pain, OR (2) any severe or horrible pain, in the 5 days prior to the MDS assessment (OBRA, PPS, or discharge) during the selected quarter.

A. Background for This QM

Unrelieved pain is a source of adverse psychological and physiological effects including nausea, dyspnea, impaired immune response, anxiety, and depression among others (Hanson, 1997; Sachs et al., 2004; Scherder & Bouma, 2000; Wrede-Seaman, 2001). Failure to identify pain can result in lower quality of life for nursing home residents—both because it can be a symptom of an underlying medical problem, and because untreated pain can result in lowered

levels of functional ability. This is especially the case in residents with cognitive impairment. Unrelieved pain costs millions of dollars annually as a result of longer hospital stays, increased rehospitalizations, increased utilization of outpatient care, and more emergency room visits (Grant et al., 1995; Wu et al., 2003). Pain management is of great interest to the public and is also one of the major goals of the Advancing Excellence in America's Nursing Homes initiative. Pain management in nursing homes is central to the Omnibus Budget Reconciliation Act of 1987 (OBRA '87) mandate to promote the "maximum practicable functioning" of residents. Failure to identify and address pain denies residents the right to freedom from neglect (Wiener et al., 2007). Evidence suggests that at least 40 percent to 85 percent of nursing home residents have persistent pain, and that pain is often not fully documented; therefore, the percentage may be even higher (Centers for Medicare & Medicaid Services, 2008; Ferrell et al., 1990; Mor et al., 2004; Parmelee et al., 1993; Sengstaken & King, 1993; Weiner & Rudy, 2002; Wu et al., 2003). A standard measure to reflect the quality of care related to pain assessment and treatment provides a benchmark for pain management practices that vary widely across nursing homes.

The MDS 3.0 pain QMs are now based on patient interview. However, one concern with the self-reporting of pain is that it may be underrepresented in patients with impaired cognitive status. One study of nursing home resident use of analgesics for pain relief found that the receipt of both opioid and nonopioid analgesics was significantly associated with the ability to self-report pain, and that residents who received analgesics had higher scores on the Mini Mental Status Exam (MMSE) and therefore higher cognitive functioning than those who were prescribed as-needed analgesics that were never administered (Allen et al., 2003). The study also found that opioid analgesic usage was associated with more speech by others to the resident (Allen et al., 2003). Analysis of the patients' medical records showed that although 86 percent of patients had been diagnosed with a painful condition, only 70 percent received analgesics. Based on this, along with the results of the study, the authors concluded that residents with impaired cognitive status were in fact less able to self-report pain and less likely to receive analgesics for pain relief (Allen et al., 2003). Methods for identifying pain in patients with cognitive impairment include the use of scales and assessments to determine the presence of symptoms such as behavioral problems or signs of depression (Leone et al., 2009); or behaviors such as crying, yelling, tactility, wincing, restricted movement, and others (Closs, 2004). Fortunately, the MDS 3.0 items for identifying cognitive impairment, BIMS was tested and showed a sensitivity of 69.7 percent to 94.4 percent, and a specificity of 85.6 percent to 78.6 percent for severe cognitive impairment, indicating that the majority of patients with cognitive impairments are being correctly identified (Chodosh et al., 2008). The BIMS showed high reliability as well: facility nurses and research assistants obtained the identical mean scores (Chodosh et al., 2008). Cognitive impairment is used as a covariate to risk adjust this QM. The risk-adjustment strategy for this QM is discussed in greater detail in Section 4.6.5.

B. Summary of Differences between MDS 2.0 and MDS 3.0 QM Definitions

The change to interview-based items was made because research has demonstrated that self-reported severity and frequency of pain using standardized scales is significantly more accurate than staff assessment of pain. QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) reports the percentage of long-stay residents who report almost constant or frequent pain and at least one episode of moderate to severe pain, OR any severe or horrible pain, in the 5 days prior to the MDS assessment (OBRA, PPS, or discharge)

during the selected quarter. The numerator is the number of long-stay residents with an MDS assessment (OBRA, PPS, or discharge) during the selected quarter and who self-report (J0200=1) almost constant or frequent pain on a scale of 1 to 4 (J0400 =1 or 2) AND at least one episode of moderate to severe pain (item J0600A = 5, 6, 7, 8, or 9 on a scale of 1–10, with 10 being the worst pain you can imagine, OR item J0600B = 2 or 3 on a scale of 0–4, with 4 being very severe, horrible pain) OR very severe/horrible pain of any frequency (item J0600A = 10 on a scale of 1 to 10 OR item J0600B = 4 on a scale of 0–4) in the 5 days prior to the assessment. The denominator is the total of all long-stay residents in the nursing facility who have an OBRA, PPS, or discharge MDS assessment during the selected quarter and who do not meet the exclusion criteria.

Target assessments are excluded if they are an admission assessment, a PPS 5-day assessment, or a PPS readmission/return assessment (A0310A=[01] or A0310B=[01,06]). Assessments are also excluded if the resident is not included in the numerator (the resident did not meet the pain symptom conditions for the numerator) AND any of the following conditions are true:

- The pain assessment interview was not completed (missing or skipped, J0200=[0,-,^]).
- The pain presence item was not completed (unknown, missing, or skippedJ0300=[9,-,^]).
- For patients with pain or hurting at any time in the last 5 days (J0300 = [1]), any of the following are true: (a) The pain frequency item was not completed (J0400=[9,-,^]); (b) neither of the pain intensity items was completed (J0600A=[99,^, -] and J0600B=[9,^,-]); or (c) the numeric pain intensity item indicates no pain (J0600A=[00]).

If the facility sample includes fewer than 30 residents, then the facility is excluded from public reporting owing to small sample size. Resident-level limited covariate risk adjustment is used in this measure. The measure adjusts for independence or modified independence in daily decision-making coded on the previous MDS assessment, as measured either by Item C1000 (=0 or 1) or C500 (=13,14, or15).

This measure was modified from the MDS 2.0 QM in several ways to address shortcomings identified by the University of Colorado as part of the DMINHo Project (Brega et al., 2007); (Brega, Goodrich, Nuccio, et al., 2008). They found that although the measure could be reported for 85.6 percent of facilities and showed a good degree of variability, the measure performed poorly on indicators of risk-adjustment adequacy and stability. From quarter to quarter, 16.5 percent of facilities experienced large (three decile or more) changes in facility ranking. However, testing of the statistical model used to risk adjust the measure showed little predictive power.

The change from the staff assessment items on MDS 2.0 to interview-based items on MDS 3.0 was made because research has demonstrated that self-reported severity and frequency of pain using standardized scales is significantly more accurate than staff assessment of pain. In addition, research showed that the MDS 2.0 measure underestimated the prevalence of pain (Saliba & Buchanan, 2008). An interview-based measure aligns pain assessment with accepted care standards across settings. As a result of going to an interview-based measure, the items used to calculate the measure and covariates changed. Specifically, the covariate for this QM can be based on either the measure of cognitive skills for daily decision making (C1000), or the

BIMS, another set of interview items (C500). This measure has resident-level limited covariate risk adjustment for residents with independence or modified independence in daily decision making as indicated on item C1000 on the prior MDS assessment or with a score of 13, 14, or 15 on C500. Based on the MDS Crosswalk Database produced by RAND, the Colorado team estimated triggering rates for this measure based on MDS 2.0 and 3.0 to be 7.7 percent and 17.5 percent, respectively, indicating that the measure would address the vast underestimation of the prevalence of pain shown by the MDS 2.0 measure.

The overall sample definitions also changed from MDS 2.0 to MDS 3.0. For the MDS 2.0, residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments. For the MDS 3.0, the analogous sample is for long-stay residents, defined as residents with more than 100 cumulative days in the facility, and assessments may be discharge, 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments, or admission, quarterly, annual, significant change, or significant correction OBRA assessments. (Some of these assessments might be excluded from the measure calculation depending on the measure.)

C. Summary of Analyses

The following sections summarize the basic analyses performed for this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Four general areas are addressed:

- number of assessments that are included and excluded from the numerator of the QM based on sample restrictions;
- whole sample prevalence of items that are related to the QM;
- findings regarding variability, reportability, reliability, and validity; and
- risk-adjustment analyses.

4.6.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 864,555 (71.8 percent) assessments included in the denominator. A total of 339,171 resident episodes (28.2 percent of the total number of long-stay resident episodes) were excluded from the denominator of QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay). ***Table 4.6-1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criteria. Note that one resident episode can meet multiple exclusion criteria. The number and percentage of resident episodes that met each exclusion criterion are as follows:

- A total of 51,303 (4.3 percent) assessments met exclusion criterion 1: missing data on items used to construct the numerator;
- 272,010 (22.6 percent) assessments met exclusion criterion 2: unable to participate in pain interview; and
- 22,402 (1.9 percent) assessments met exclusion criterion 3: the target assessment was an admission, a PPS 5-Day assessment, or a PPS readmission/return assessment.

Numerator Items – Distribution of Triggering Events

Table 4.6-2 shows the frequency and percentage of triggering events (items that can trigger the numerator of the QM). J0400 asks about the frequency of pain in the last 5 days. The response categories are 1 (almost constantly) to 4 (rarely). Item J0600A asks about pain intensity using a 0 (no pain) to 10 (worst imaginable pain) scale. J0600B asks about pain using a verbal descriptor scale with a range of 1 (mild pain) to 4 (very severe, horrible). For all three items, a code of 9 indicates that the resident was unable to answer. Overall, the triggering events were distributed across all scores, and don't show a floor or ceiling effect. The table shows that the frequency of responses to item J0400 is greatest for responses of 2 and 3, indicating that the majority of residents triggering the QM report frequent (2) or occasional (3) pain in the past 5 days. On Item J0600A, the highest frequencies were seen for responses of 4, 5, and 6 (corresponding to moderate pain) on a 10-point pain scale. Similarly, on J0600B, the highest frequency was seen for response 2 (moderate pain). The pattern for short-stay pain measure, QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay), is similar, though the percentages of residents reporting pain at any level are notably lower in the long-stay sample than in the short-stay sample. This may be in part because long-stay residents are less likely to be receiving rehabilitation services such as physical therapy for conditions such as hip or knee replacements where pain would be expected, even with a regularly scheduled pain regimen in place. Also notable is the frequency of carets for each pain item. Carets indicate a skip based on response to Item J0300 ("have you had pain or hurting at any time in the last 5 days?"). Items J0400, J0600A, and J0600B are skipped for roughly 75 percent, 82 percent, and 88 percent of residents, respectively. These numbers are higher than for the short-stay pain measure, which is consistent with the findings that the prevalence of pain is lower among long-stay residents and/or that the prevalence of residents unable to answer the J0300 ("have you had pain or hurting at any time in the last 5 days?") is higher among long-stay residents.

Table 4.6-1
Target assessments used in the calculation of QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	864,555	71.8%
Excluded—Exclusion 1	51,303	4.3%
Excluded—Exclusion 2	272,010	22.6%
Excluded—Exclusion 3	22,402	1.9%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Missing data

Exclusion 2 = Unable to participate in pain interview

Exclusion 3 = Wrong type of assessment (admission, PPS 5-day assessment, or a PPS readmission/return assessment)

[a] Column values may not add up to total because a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Table 4.6-2
Responses to items important to calculating numerator for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

MDS 3.0 Item	Entry	Frequency	Percentage	Cumulative %
J0400	— = missing	34,257	2.9%	2.9
J0400	1 = Almost constantly	33,660	2.8%	5.6
J0400	2 = Frequently	84,436	7.0%	12.7
J0400	3 = Occasionally	129,011	10.7%	23.4
J0400	4 = Rarely	24,055	2.0%	25.4
J0400	9 = Unable to answer	6,734	0.6%	25.9
J0400	^ (skipped)	891,573	74.1%	100.0
J0400	Total	1,203,726	—	—
J0600a	— (missing)	39,454	3.3%	3.3
J0600a	0	331	0.0%	3.3
J0600a	1	2,227	0.2%	3.5
J0600a	2	10,184	0.8%	4.3
J0600a	3	19,051	1.6%	5.9
J0600a	4	23,777	2.0%	7.9
J0600a	5	34,082	2.8%	10.7
J0600a	6	25,890	2.2%	12.9
J0600a	7	20,126	1.7%	14.5
J0600a	8	25,874	2.2%	16.7
J0600a	9	7,319	0.6%	17.3
J0600a	10	11,619	1.0%	18.3
J0600a	99	11,083	0.9%	19.2
J0600a	^ (skipped)	972,709	80.8%	100.0
J0600a	Total	1,203,726	—	—
J0600b	— (missing)	41,784	3.47%	3.47
J0600b	1 = Mild	35,353	2.94%	6.41
J0600b	2 = Moderate	62,553	5.2%	11.6
J0600b	3 = Severe	17,180	1.43%	13.03
J0600b	4 = Very severe, horrible	2,469	0.21%	13.24
J0600b	9 = Unable to answer	5,624	0.47%	13.7
J0600b	^ (skipped)	1,038,763	86.3%	100.0
J0600b	Total	1,203,726	—	—

NOTE: J0400 = Pain Frequency; J0600a = Numeric Rating Scale; J0600b = Verbal Descriptor Scale

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011a (\\quarter_4_5\\nh_011_10.log)

4.6.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with perfect scores (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes), which may indicate a "ceiling effect." ***Table 4.6-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) is 11.5 percent with a standard deviation of 8.5 percent. This QM has an interquartile range of 11.5 percent. About 5.9 percent of facilities have scores of 0 percent (perfect scores).

Table 4.6-3
QM score distribution for QM #0677 Percentage of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

Risk adjustment used	*n*	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with perfect scores	Interquartile range
Unadjusted	12,059	12.3%	9.1%	2.0%	5.3%	10.6%	17.8%	25.0%	5.6%	12.5%
Adjusted	11,896	11.5%	8.5%	1.9%	5.0%	9.9%	16.5%	23.2%	5.9%	11.5%

Analysis date: 5/29/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\combined_fa_adj\nh_011_10_ra_combined_v3.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 long-stay residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We examined the percentage of nursing homes that can report each measure (referred to as the QM reportability). ***Table 4.6-4*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 78.5 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) using MDS 3.0. This is lower than the reportability for the MDS 2.0 measure, which 85.6 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008). The difference

may be due to the changes in QM specifications. The denominator for the MDS 3.0 QM includes only residents whose pain assessment interview was completed. About 22.6 percent of long-stay residents were unable to participate in a pain interview (***Table 4.6-1***).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may improve the reportability of long-stay QMs to some extent, but the impact on the long-stay QM should be less marked than the short-stay QMs.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay), we examined changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.6-4***. After excluding discharge assessments from the target assessment sample, reportability is slightly increased for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) (75.8 percent with discharge assessments vs. 75.7 percent without discharge assessments). This increase may be due to the higher rate of missing value for items used to create this QM on discharge assessments.

Table 4.6-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator	864,555 (71.8%) [a]	884,453 (73.5%) [a]
Number (%) of resident episodes in the numerator	101,684 (8.4%) [a]	106,832 (8.9%) [a]
Number (%) of facilities able to report this QM	11,896 (75.8%) [b]	11,879 (75.7%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay sample (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.6-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. For each pair of quarters, large changes in ranking were rare. About two-thirds of facilities stayed in approximately the same rank (the change was within one decile), and relatively few (about 7 percent) facilities had a rank change by more than three deciles.

Table 4.6-5
Distribution of change in facility rank from one quarter to the next, QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

Magnitude of Shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	6,748 (63.1%)	7,125 (63.6%)	7,441 (65.3%)
Between 1 and 2 deciles	2,086 (19.5%)	2,221 (19.8%)	2,151 (18.9%)
Between 2 and 3 deciles	1,120 (10.5%)	1,038 (9.3%)	1,066 (9.4%)
More than 3 deciles	748 (7.0%)	812 (7.3%)	740 (6.5%)
Total	10,702	11,196	11,398

NOTES:

Total number of facilities in each column reflects all the facilities that could report the measure in both quarters.

Columns may not total 100.0% due to rounding.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 1 through Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.6-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 8.5 percent). Overall, the average change in scores for this measure was a decline of 0.5 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (73.6 percent saw changes of less than one standard deviation), whereas just over 1 percent saw changes of more than three standard deviations.

Table 4.6-6
Change in facility score from one quarter to the next, QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
11,398	-0.5	5.3%	0.7%	2.1%	10.5%	34.9%	38.7%	10.5%	2.2%	0.6%

NOTES:

Number of facilities reporting (11,398) reflects facilities that reported this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated). For example, the "pain group" of measures, QM #0677 [Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)] and QM #0676 [Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay)], should reflect the quality of related clinical care processes. Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures where applicable. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) and the facility's percentile rank on QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) for Quarter 4, 2011. We found that the correlation between the nursing home's percentile rank on QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) and percentile rank on QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) was moderate (correlation coefficient = .560) and statistically significant ($p <$ 0.001).[17]

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether State characteristics might be a source of facility score variation for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay), we conducted a one-way analysis of variance (ANOVA). We also examined the interquartile difference between the mean State-level scores across States. The proportion of variance in this measure explained by the State in which facilities are located is 9.5 percent and significant [$F(50, 11{,}845) = 24.9$, $p < .001$].[18] The difference between the mean State-level scores for States at the 25th percentile and the 75th percentile is 6.4 percentage points. Thus, although the majority of the variance in QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) scores is due to factors other than geography, a small and significant proportion of the variance is explained by the respective States in which nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns of missing data indicate that certain types of residents tend to have assessments with missing data in ways that affect the calculation of a quality measure, then that

---

[17] RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log).

[18] RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request).

measure may be not be capturing processes and outcomes of care for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

In Quarter 4 of 2011, there were 339,171 residents (28.2 percent of the long-stay population) excluded from the construction of this model-based risk-adjusted quality measure. There were 22,402 (1.9 percent of the population) that were excluded because the target assessment was an admission, PPS 5-day assessment, or readmission/return assessment. An additional 272,010 (22.6 percent) were excluded because residents could not participate in the pain assessment interview, meaning that they were rarely or never understood (J0200 = [0]) or they were unable to answer when asked "Have you had pain or hurting at any time in the last 5 days?" (J0300 = [9]). Out of the remaining residents, 51,303 (4.3 percent) were excluded due to missing data on either target assessments or on prior assessments (which are used to calculate the covariate).

For this measure, we analyzed whether missing data on pain items J0300, J0400, or J0600 varied systematically on the following characteristics: age greater than or equal to 85, sex, a score on the Brief Interview for Mental Status (BIMS) less than or equal to 12 (scores indicating cognitive impairment), and inability to make oneself understood (from item B0700). ***Table 4.6-7*** summarizes the results of this analysis. For this measure, age and sex appear to have only minimal effects on rates of missing data for pain items. The inability to make oneself understood and cognitive impairment (which are likely related) appear to have effects: residents with missing data on these items are more than five times as likely to have problems with communication (1.1 percent vs. 5.8 percent) and somewhat more likely to be cognitively impaired (60.5 percent vs. 72.4 percent). Because this measure relies on the ability of a resident to report his or her own pain, an inability to make oneself understood might naturally lead to difficulty for a facility in properly completing these items on the target assessment.

Table 4.6-8 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. The threat posed to validity by missing data on items used to calculate this measure appears small, with a mean missing data rate of 4.5 percent and a median rate of 2.2 percent, but there appear to be outlying facilities with higher rates. It may be instructive to examine the characteristics of these facilities to see whether the higher missing rates are associated with a systematic bias. Future analyses of this issue need to consider using supplemental data sources (e.g., the Provider of Services file) for facility characteristics.

We further examined the possible association between missing data and QM scores. ***Table 4.6-9*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure. The mean facility-level missing rate monotonically increases as a function of quartile, from the quartile with the lowest scores (missing rate= 3.7 percent) to that with the highest scores (5.3 percent). This pattern is also shown by the significant but small correlation between missing data and QM scores for this measure ($r = .107$, $p < .001$). Thus, facilities with higher levels of missing data rates also tend to have greater proportions of residents who self-report moderate to severe pain.

Table 4.6-7
Missing data compared across selected resident characteristics

Resident characteristics	Any missing data on J0300, J0400, or J0600 *n*	Any missing data on J0300, J0400, or J0600 % [a]	No missing data on J0300, J0400, or J0600 *n*	No missing data on J0300, J0400, or J0600 % [b]
Age ≥ 85	19,603	39.2%	378,018	43.7%
Male	17,311	34.6%	280,324	32.4%
BIMS <= 12	11,475	72.4%	452,384	60.5%
Inability to make oneself understood (B0700 = 3)	2,459	5.8%	9,611	1.1%

Analysis date: 6/28/2012

BIMS = Brief Interview for Mental Status.

NOTES:

[a] Percentage reflects proportion of all target assessments with any missing data on J0300, J0400, or J0600 and no missing data on BIMS or B0700.

[b] Percentage reflects proportion of all target assessments with no missing data on J0300, J0400, or J0600 and no missing data on BIMS or B0700.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db157_request\db157_request.log)

Table 4.6-8
Distribution of facility-level missing rate for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,399	4.5%	7.0%	0.0%	0.0%	2.2%	6.3%	11.4%

NOTES:

n = number of facilities that have data for numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet minimum sample size for reporting.

Analysis date: 6/28/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db158_request\db158_request.log)

Table 4.6-9
Distribution of facility-level missing rate for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) by quartile of QM score

Quartile of QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,827	3.7%	6.9%	0.0%	0.0%	1.3%	4.8%	10.0%
26%–50%	3,827	4.0%	5.8%	0.0%	0.0%	2.1%	5.7%	10.3%
51%–75%	3,829	4.5%	6.0%	0.0%	0.0%	2.6%	6.6%	11.5%
76%–100%	3,794	5.4%	7.1%	0.0%	0.0%	3.2%	7.7%	13.2%

NOTES:

Total *n* = number of facilities that have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in scores after excluding discharge assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2), the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.6-10.*** The mean nursing home–level QM score and distributions change slightly after excluding discharge assessments from the set of eligible target assessments. The average facility score increased by 0.3 percent (with a median change of 0 percent) when discharge assessments were excluded. Thus, the use of the discharge assessment in the QM score calculations does not meaningfully alter the distribution of scores among facilities.

Table 4.6-10
Distribution of facility-level score change after excluding discharge assessment for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
11,879	0.3%	1.5%	-1.1%	-0.3%	0%	1.0%	2.2%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

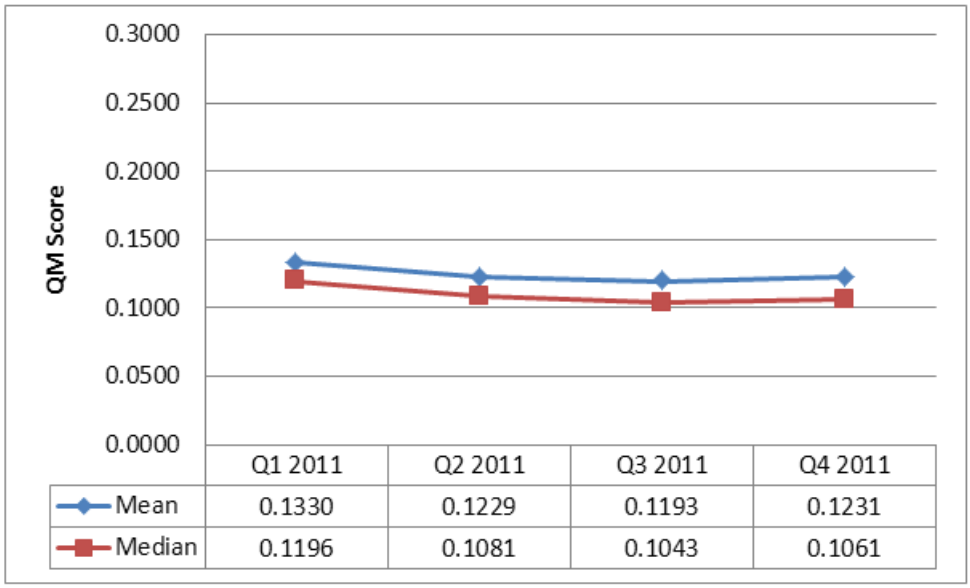
SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request_v1\db135_request_v1.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address this interest in whether seasonal variation might play a role in the score for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay), we examine the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.6-1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, was 13.3 percent, 12.3 percent, 11.9 percent, and 12.3 percent. The quarterly national median was 12.0 percent, 10.8 percent, 10.4 percent, and 10.6 percent. Both mean and median fell between Quarter 1, 2011, and Quarter 3, 2011, and rose slightly in Quarter 4, 2011. With only 12 months of data available, it is too early to ascertain whether the changes in national QM score over time are consistent with seasonal variation, this will need to be confirmed when multiple years of data become available.

Figure 4.6-1
Seasonal (quarterly) variation in QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)



	Q1 2011	Q2 2011	Q3 2011	Q4 2011
Mean	0.1330	0.1229	0.1193	0.1231
Median	0.1196	0.1081	0.1043	0.1061

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 to Quarter 4, 2011 (\qm_quarter_1_2\complete\nh_011_10.log \qm_quarter_2_3\ complete\nh_011_10.log \qm_quarter_3_4\ complete\nh_011_10.log \qm_quarter_4_5\ complete\nh_011_10.log)

4.6.5 Risk Adjustment

QM #0677 Percent of Residents who Self-Report Moderate to Severe Pain (Long Stay) uses limited resident-level covariate risk adjustment in addition to exclusions. The measure is restricted to residents who can self-report pain, excluding a set of residents at risk for

underreporting of pain. For the covariate risk adjustment (indirect standardization), the measure is adjusted for independence or modified independence in daily decision making (C1000=0 or 1) for residents evaluated with the staff assessment, or no cognitive impairments (C0500>12) coded on the previous MDS assessment. The frequency of the covariate is shown in ***Table 4.6-11*** below. The right-most column of Table 4.6-11 shows that 36.1 percent of the sample are either moderately to severely impaired in daily decision making or have a BIMS score lower than 12.

To understand how facilities' QM scores may be influenced by model-based risk adjustment, we present an analysis below (Table 4.6-11) that reports mean risk-adjusted QM scores compared with unadjusted QM score. For facilities with 30 or more residents, the minimum facility sample size for public reporting, the median percentage of long-stay residents self-reporting moderate to severe pain is 10.8 percent when this QM is not adjusted and 10.7 percent when it is fully adjusted. The proportion of facilities that have a significantly greater risk-adjusted QM score (defined as more than 5 percent of the adjusted mean score) is slightly smaller than the proportion of facilities whose scores decline significantly (27.0 percent vs. 31.7 percent). The results indicate that although national mean QM score did not change substantially before and after risk adjustment, different facilities are subject to changes of different sizes and directions (increase vs. decrease).

Table 4.6-11
Frequency of covariate used for risk adjustment:
QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

Covariate name: Independence or modified independence in daily decision making	Facility-level: unadjusted-adjusted QM	Facility-level: fully adjusted QM	# of facilities that improve by more than 5% of adjusted mean	# of facilities that decline by more than 5% of adjusted mean	Assessment level: # assessments: covariate = 1	Assessment level: % assessments: covariate = 1
Covariate1=1 (residents from all facilities)[1,2]	—	—	—	—	310,063	36.1%
Facilities with sample>=30						
# facilities	12,168	11,961	3,235	3,792	—	—
QM median	10.8%	10.7%	—	—	—	—

NOTES:

1. The sample for this frequency was the initial assessments used to calculate risk-adjusted QM score (n = 858,524).

2. Covariate definition:

Independence or modified independence in daily decision making (C1000=0 or 1) for residents evaluated with the staff assessment; or no cognitive impairments (C0500>12).

Analysis date: 7/17/12

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db180_171_169\db180_request.log)

To examine whether the performance of the risk-adjustment model for this QM is stable across quarters, we compared the intercepts and coefficients for the covariates for each quarter in 2011. In ***Table 4.6-12***, we show that values for the intercept are very similar across quarters. Coefficients for the covariate increased slightly across quarters, suggesting that the difference in

the possibility of self-reporting pain becomes bigger over time, although at a very slow pace. The model goodness-of-fit statistic-Pseudo R-square also has comparable values across four quarters. Overall, these results suggest that the performance of the risk adjustment is stable in 2011 and thus support the validity of the risk-adjustment model.

Table 4.6-12
Intercepts and coefficients by quarter of data for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)

Covariate in 2011	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Intercept	-2.369	-2.447	-2.481	-2.525
Independence or modified independence in daily decision making	1.017	1.053	1.069	1.091
Pseudo R-square	0.036	0.039	0.040	0.042

NOTES:

Independence or modified independence in daily decision making on the prior assessment:

Covariate = 1 if C1000 = [0, 1] or if (C0500 ≥ [13] and C0500 ≤ [15])

Covariate = 0 if C1000 = [2, 3] or if (C0500 ≥ [00] and C0500 ≤ [12]).

Covariate = missing if either of the following is true:

1. C0500 = [99,−,^] and C1000 = [−,^].
2. No prior assessment is available.

Analysis date:

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 to Quarter 4, 2011 (\quarter_4_5\db170_request\db170_request.log)

To understand how model specification influences the performance of the risk-adjustment model and the adjusted QM scores, we estimated the risk-adjustment model for this QM using both a single-level model and a hierarchical model. The current specification of the risk-adjustment model for this QM uses single-level logistic regression, which does not take into account the fact that residents are clustered within nursing homes. The hierarchical model, on the other hand, addressed the issue. We estimated the hierarchical model using the specification of multilevel mixed-effects logistic regression. ***Table 4.6-13*** shows that coefficients for the covariates from the single-level model and the ones from the hierarchical model do not differ substantially (1.091 vs. 1.141), but the hierarchical model has a much higher predictive power. The C-statistics indicate that the single-level model has acceptable power for predicting self-report pain (C-stat = 0.632), and the hierarchical model has nearly excellent predictive power for this QM (C-stat = 0.771).

The risk-adjusted QM score based on single-level model differs from the ones based on the hierarchical model. For an average facility, the risk-adjusted QM score based on the single-level model is slightly higher than the one based on the hierarchical model (difference = 1.3 percentage point). The interquartile range for facility-level difference in risk-adjusted QM score between single-level model and hierarchical model is -3.4 percent to 4.3 percent. Based on these

risk-adjusted QM scores, we also compared facility rank, with a particular interest in facilities between 95 and 100 percentile (indicating extremely poor quality of care). Among the 444 facilities identified by the single-level model as poor-quality outliers, only 231 (52.0 percent) were identified as poor-quality outliers based on the hierarchical model.[19]

These results suggest that the risk-adjustment model for this QM performs differently with different model specifications. As a result, the risk-adjusted QM scores and facility rank based on these scores show differences.

Table 4.6-13
Intercepts and coefficients by quarter of data for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay)—Comparison between single-level model and hierarchical generalized linear model

Covariate in Quarter 4, 2011	Single-level model covariate	Single-level model 95% confidence interval	Hierarchical model covariate	Hierarchical model 95% confidence interval
Independence or modified independence in daily decision making	1.091	1.077 – 1.104	1.141	1.126 – 1.155
C-statistic	0.632	—	0.771	—

NOTES:

Independence or modified independence in daily decision making on the prior assessment:

Covariate = 1 if C1000 = [0, 1] or if (C0500 ≥ [13] and C0500 ≤ [15])

Covariate = 0 if C1000 = [2, 3] or if (C0500 ≥ [00] and C0500 ≤ [12]).

Covariate = missing if either of the following is true:

1. C0500 = [99,−,^] and C1000 = [−,^].
2. No prior assessment is available.

Analysis date: 6/1/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db130_request\nh_011_10_xtra_state.log)4.7 QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

4.7 QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

4.7.1 Summary of Findings

QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) reports the percentage of long-stay residents at a high risk for pressure ulcers who have Stage 2 to 4 pressure ulcers. This QM is risk adjusted using denominator exclusions. Residents not at a high risk for

[19] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db130_request\db130_request.log)

pressure ulcers—defined as impaired bed mobility or transfer, comatose or malnutrition or at risk of malnutrition—are excluded from the denominator.

We found that 67.8 percent of the long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 72.4 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 6.9 percent. The QM score varies across facility with a standard deviation of 4.6 percent and interquartile range of 6.2 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To examine the reliability of this QM, we analyzed the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. About half of facilities remained within the same decile from quarter to quarter. About 17 percent of facilities have rank changes of more than three deciles from quarter to quarter. Compared with other QMs, the relatively large rank changes may be due to the restrictive denominator inclusion criteria (i.e., high risk) and thus smaller denominator size.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a care process group. The hypothesis was that changes in one QM within a measure group (e.g., short-stay and long-stay pressure ulcers QMs) should be correlated with changes in related measures because they reflect similar care processes. Findings from this analysis showed significant but weak correlations among the short- and long-stay pressure ulcer measures (correlation $r = 0.148$). The weak correlation may be due to the different specifications for the short-stay and long-stay QMs (i.e., the short-stay QM is an incidence measure whereas the long-stay QM is a prevalence measure). It may also suggest that the care processes for preventing pressure ulcers and/or the patient characteristics in the short-stay versus long-stay populations are very different, and that facilities may have more specialized expertise in dealing with only one of the patient groups. Second, we examined the frequency and distribution of missing data. The rate of missing data is very low (nearly zero) for items used to calculate this QM and therefore should not pose a threat to validity of the QM. Finally, we analyzed the potential geographic (State) and seasonal (quarter) variations in this QM. The State in which the facility is located explains about 6.5 percent of the variation in this QM. The QM score remained largely stable from Quarter 1 to 4 in 2011.

Because reliability and validity tests support acceptable reliability and validity of this QM, QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) received full NQF endorsement on August 1, 2012.

4.7.2 Background and Introduction to QM

QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) reports the percentage of all long-stay residents who were identified as at high risk for pressure ulcers who have one or more Stage 2 to 4 pressure ulcer(s) on their target assessments. High-risk populations are those who are comatose, have impaired bed mobility or transfers, or are suffering

from malnutrition. This measure excludes Stage 1 ulcers from the definition. Long-stay residents are those who have been in nursing facility care for more than 100 days. Target assessments may be discharge; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

A. Background for This QM

Pressure ulcers are serious medical conditions. They typically result from prolonged periods of uninterrupted pressure on the skin, soft tissue, muscle, and bone (Bates-Jensen, 2001; Institute for Healthcare Improvement, 2007; Russo et al., 2008). Vulnerable patients include the elderly; patients with stroke or diabetes, dementia, circulatory diseases, dehydration, malnutrition, frailty, and feeding tubes; and people who use wheelchairs or are bedridden—that is, any patient with impaired mobility or sensation (Bates-Jensen, 2001; Gumieiro et al., 2012; Hurd, T., et al., 2010; MacLean, 2003; Teno et al., 2012). Pressure ulcers interfere with the activities of daily living, predispose patients to osteomyelitis and septicemia, and are strongly associated with longer hospital stays and mortality (Bates-Jensen, 2001).

Pressure ulcers are high-volume and high-cost adverse events across the spectrum of health care settings from acute hospitals to home health (Russo, Steiner, & Spector, 2006); (Bates-Jensen, 2001; Hurd, T., et al., 2010). The prevalence of pressure ulcers in health care facilities is increasing: about 2.5 million patients are treated annually for pressure ulcers in acute-care facilities (Institute for Healthcare Improvement, 2007; Russo et al., 2006). In 2006, there were 503,300 acute-hospital stays during which pressure ulcers were noted—a 78.9 percent increase from 1993 when there were about 281,300 hospital stays related to pressure ulcers (MacLean, 2003; Russo et al., 2006). Additionally, approximately 60,000 patients die from pressure ulcer complications annually (Institute for Healthcare Improvement, 2011).

Pressure-ulcer incidence rates vary considerably by clinical setting—in 2007 numbers ranged from 0.4 percent to 38 percent in acute care, from 2.2 percent to 23.9 percent in skilled nursing facilities and nursing homes, and from 0 percent to 17 percent in home care (Duncan, 2007; Institute for Healthcare Improvement, 2007). In 2009, the incidence of pressure ulcers at Stage 2 or higher in nursing homes was 2.4 per 100 residents, and the prevalence was 7.3 per 100 residents (Centers for Medicare & Medicaid Services, 2010a). Incidence and prevalence are related to a facility's patient acuity, but also to its prevention practices; initiatives to improve pressure-ulcer prevention can reduce incidence and prevalence, with one 4-year prevention campaign reducing the incidence from 5.19 percent to 0.73 percent (Tippet, 2009).

Patients with acute-care hospitalizations related to pressure ulcers were more likely to be discharged to long-term care facilities (e.g., a skilled nursing facility, an intermediate care facility, or a nursing home), than were hospitalizations for all other conditions (Hurd, T., et al., 2010; Institute for Healthcare Improvement, 2007). In fact, more than half of principal pressure ulcer stays (53.4 percent) and secondary pressure ulcer stays (54.5 percent) were discharged to long-term care: more than 3 times the rate of hospitalizations for all other conditions (16.2 percent) (Hurd, T., et al., 2010).

As reported in the 2004 National Nursing Home Survey results, about 159,000 current U.S. nursing home residents (11 percent) had pressure ulcers. Stage 2 ulcers were the most

common, accounting for about 50 percent of all pressure ulcers. Stages 1, 3, and 4 made up about the other 50 percent of all ulcers (Park-Lee & Caffrey, 2009). Stage 1 pressure ulcers are not included in the proposed quality measure; researchers have suggested that inclusion of Stage 1 pressure ulcers in the quality measures adds little value (Brega, Goodrich, Nuccio, et al., 2008; Lynn et al., 2007).

Previous research examined pressure-ulcer incidence and prevalence across post-acute settings (Hurd, Moore, et al., 2010). The prevalence and incidence of pressure ulcers was examined across post- acute settings using MDS 2.0 data from April 1, 2006, through July 15, 2006 (Abt Associates Inc. et al., 2007). The prevalence of pressure ulcers Stages 1 to 4 was 13 percent, and the prevalence of Stages 3 to 4 ulcers was 3 percent nationwide (Hurd, T., et al., 2010). Pressure ulcers may cause extreme discomfort to the patient and often lead to serious, life-threatening infections, which substantially increase the total cost of care (Agency for Healthcare Research and Quality, January 2009; National Pressure Ulcer Advisory Panel Board of Directors, 2001; Russo et al., 2006). The main driver of cost is the presence of complications, which involve diagnostic tests, additional monitoring, more expensive pressure-relieving surfaces, and extended lengths of stay (Agency for Healthcare Research and Quality, January 2009).

In 2006, the Federal Register reported 322,946 cases of Medicare patients with a pressure ulcer as a secondary diagnosis—each case had an average charge of $40,381 for a hospital stay, for an annual total cost of $13 billion (Centers for Medicare & Medicaid Services, 2007). In 2008, there were 394,699 cases of pressure ulcers in Medicare patients, with the cost of treating the pressure ulcer alone at $8,730 per case, resulting in a total cost of more than $3.4 billion for the pressure ulcer treatment (Van Den Bos et al., 2011). The Advancing Excellence in America's Nursing Homes Campaign, a national effort launched in 2006 to help nursing homes measurably to improve care, reported that it can cost as much as $19,000 to treat a single Stage 4 pressure ulcer.

B. Summary of Differences between MDS 2.0 and MDS 3.0 QM Definitions

The underlying data items used to calculate the MDS 3.0 QM were revised to include items or language used in other health care settings in order to improve communication across settings and providers (e.g., the pressure ulcer items included in the National Pressure Ulcer Advisory Panel's PUSH tool are used to describe pressure ulcers in the MDS 3.0) (Saliba & Buchanan, 2008). Therefore, the proposed measure based on the new MDS 3.0 pressure ulcer items better aligns the measure with accepted best practices.

The MDS 3.0:

- Eliminates reverse staging, which does not reflect the pathophysiology of pressure-ulcer healing
- Pressure ulcer staging based on deepest anatomical change (recommendation Wound, Ostomy, and Continence Nurses Society (WOCN), National Pressure Ulcer Advisory Panel (NPUAP)
- Unstageable pressure ulcers are assessed as separate items (NPUAP, WOCN)

- The number of pressure ulcers that were present on admission is collected for each stage
- Tissue type for most advanced stage
- Definitions are based on NPUAP
- Unstageable pressure ulcers are assessed as separate items (NPUAP, WOCN)

Stage 1 pressure ulcers are not included in the proposed quality measure; researchers have suggested that inclusion of Stage 1 pressure ulcers in the quality measures adds little value, penalizes facilities for early identification, and they are inconsistently assessed especially for populations with darker skins.

Note also that the overall sample definitions changed from MDS 2.0 to MDS 3.0. For the MDS 2.0, residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments and typically capture residents with a 30-day average length of stay. For the MDS 3.0, the analogous sample is for long-stay residents, defined as residents with more than 100 cumulative days in a facility, and assessments may be discharge; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections summarize the basic analyses performed for this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- number of assessments that are included and excluded from the numerator of the QM based on sample restrictions;
- whole sample prevalence of items that are related to the QM; and
- findings regarding variability, reportability, reliability, and validity.

4.7.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 816,508 (67.8 percent) assessments included in the denominator. A total of 18,799 resident episodes (2.3 percent of the total number of long-stay resident episodes) were excluded from the denominator of QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay). ***Table 4.7-1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criteria. Note that one resident episode can meet multiple exclusion criteria. The number and percentage of resident episodes that met each exclusion criterion are as follows:

- A total of 699 (0.0 percent) assessments met exclusion criterion 1: missing data on items that give pressure ulcer counts;

- 18,102 (2.2 percent) assessments met exclusion criterion 2: the target assessment was an admission assessment or a PPS 5-day or readmission/return assessment; and
- 387, 218 (32.3 percent) met exclusion criterion 3: lacking the characteristics for membership in the high-risk sample.

Table 4.7-1
Target assessments used in the calculation of QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	816,508	67.8%
Excluded—Exclusion 1	699	0.0%
Excluded— Exclusion 2	18,102	1.5%
Excluded— Exclusion 3	387,218	32.2%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Missing Data

Exclusion 2 = Wrong type of target assessment (admission or PPS 5-day or readmission/return assessment)

Exclusion 3 = Low Risk

[a] Column values may not add up to total because a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items – Distribution of Triggering Events

Table 4.7-2 illustrates the frequency of pressure ulcers, by Stages 2 to 4, on the target assessments used in calculating the QM. The frequencies also capture potential differences between high-risk residents and residents not classified as high risk. We hypothesized that high-risk residents would demonstrate a higher proportion of pressure ulcers, regardless of stage, compared with residents not classified as high risk. Results indicate that high-risk residents did have a higher proportion of Stage 2 to 4 pressure ulcers compared with residents not classified as high risk. For example, almost 6.0 percent of high-risk residents (48,714/816,508) had one pressure ulcer classified as Stage 2, 3, or 4. Less than 1 percent of residents not classified as high risk (3,274/363,004) had one pressure ulcer classified as Stage 2, 3, or 4. High-risk residents with one Stage 2 pressure ulcer were the most frequently reported category (3.3 percent of all high-risk residents), followed by high-risk residents with one Stage 3 ulcer (1.4 percent). Regardless of risk classification, the most common total number of pressure ulcers on the target assessment was one, indicating that although pressure ulcers are common, the majority of residents have only one.

Table 4.7-2
Pressure ulcer frequency on target assessments by stage of pressure ulcer

# of Pressure Ulcers	M0300B. Number of Stage 2 pressure ulcers not high risk *n* (%)	M0300B. Number of Stage 2 pressure ulcers high risk *n* (%)	M0300C. Number of Stage 3 pressure ulcers not high risk *n* (%)	M0300C. Number of Stage 3 pressure ulcers high risk *n* (%)	M0300D. Number of Stage 4 pressure ulcers not high risk *n* (%)	M0300D. Number of Stage 4 pressure ulcers high risk *n* (%)
0	2,911 (0.8%)	39,638 (4.9%)	4,551 (1.3%)	61,728 (7.6%)	4,659 (1.3%)	62,285 (7.6%)
1	2,021 (0.6%)	27,234 (3.3%)	699 (0.2%)	11,111 (1.4%)	554 (0.2%)	10,369 (1.3%)
2	347 (0.1%)	6,003 (0.7%)	87 (0.0%)	1,418 (0.2%)	96 (0.0%)	1,468 (0.2%)
3	58 (0.0%)	1,339 (0.2%)	19 (0.0%)	449 (0.1%)	37 (0.0%)	464 (0.1%)
4	19 (0.0%)	450 (0.1%)	3 (0.0%)	90 (0.0%)	12 (0.0%)	196 (0.0%)
5	4 (0.0%)	118 (0.0%)	0 (0.0%)	38 (0.0%)	1 (0.0%)	51 (0.0%)
6	2 (0.0%)	69 (0.0%)	1 (0.0%)	7 (0.0%)	0 (0.0%)	17 (0.0%)
7	1 (0.0%)	22 (0.0%)	0 (0.0%)	7 (0.0%)	0 (0.0%)	6 (0.0%)
8	1 (0.0%)	10 (0.0%)	0 (0.0%)	2 (0.0%)	1 (0.0%)	3 (0.0%)
9	1 (0.0%)	20 (0.0%)	0 (0.0%)	4 (0.0%)	0 (0.0%)	8 (0.0%)
No unhealed PUs	357,635 (98.52%)	741,572 (90.82%)	357,635 (98.5%)	741,572 (90.8%)	357,635 (98.5%)	741,572 (90.8%)
Total	363,004	816,508	363,004	816,508	363,004	816,508

Analysis date: 7/2/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 t (\quarter_4_5\db175_request\db175_request.log)

In ***Table 4.7-3*** we show the combinations of pressure ulcers represented among residents who trigger the numerator of this QM. Consistent with the prior table, Stage 2 ulcers are the most common reason for being counted in the numerator. Residents with only Stage 2 ulcers account for 54.0 percent of the numerator, followed by residents with only Stage 3 ulcers (15.9 percent of numerator), and residents with only Stage 4 ulcers (14.4 percent of numerator).

Table 4.7-3
Pressure-ulcer profiles for residents included in the numerator of QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

Any Stage 4	Any Stage 3	Any Stage 2	Diagnosis of pressure ulcer*	Count	Percent of numerator	Percent of denominator
0	0	1	0	30,670	54.0%	3.8%
0	0	1	1	949	1.7%	0.1%
0	1	0	0	9,017	15.9%	1.1%
0	1	0	1	641	1.1%	0.1%
0	1	1	0	1,715	3.0%	0.2%
0	1	1	1	133	0.2%	0.0%
1	0	0	0	8,203	14.4%	1.0%
1	0	0	1	1,397	2.5%	0.2%
1	0	1	0	1,182	2.1%	0.1%
1	0	1	1	180	0.3%	0.0%
1	1	0	0	1,047	1.8%	0.1%
1	1	0	1	137	0.2%	0.0%
1	1	1	0	382	0.7%	0.0%
1	1	1	1	54	0.1%	0.0%
0	0	0	1	1,079	1.9%	0.1%
0	0	0	0	759,722	N/A	93.0%
Total	—	—	—	816,508	—	—

NOTE: 0 = absent, 1 = present

*Diagnosis of pressure ulcer = 1 if MDS item I8000 indicates any of the following ICD-9 codes: 707.22, 707.23, 707.24.

Analysis date: 7/2/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db175_request\db176_request.log)

4.7.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with perfect scores (i.e., 0 percent triggering rates for

measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes) which may indicate a "ceiling effect." ***Table 4.7-4*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) is 6.9 percent with a standard deviation of 4.7 percent. This QM has an interquartile range of 6.2 percent. About 6.9 percent of facilities have scores of 0 percent (perfect scores).

Table 4.7-4
QM score distribution for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
11,352	6.9%	4.7%	1.7%	3.3%	6.3%	9.5%	12.9%	6.9%	6.2%

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting of this quality measure.

QM scores are reported at the facility-level

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_013_10.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 long-stay residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examine the percentage of nursing homes that can report each measure (referred to as the QM reportability). ***Table 4.7-5*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 72.4 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) using MDS 3.0. This is an improvement over the MDS 2.0 measure, which 62.2 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008). The differences between MDS 2.0 and MDS 3.0 discussed in Section 4.7.2 may explain this improvement.

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the

inclusion of the discharge assessment may improve the reportability of long-stay QMs to some extent. Therefore, the reportability of short-stay QMs may be substantially improved after including the discharge assessment, but the impact on the long-stay QM should be less marked.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay), we report changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.7-5***. After excluding discharge assessments from the target assessment sample, reportability is largely unchanged for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) (72.4 percent with discharge assessments vs. 71.9 percent without discharge assessments).

Table 4.7-5
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator	816,508 (67.8%) [a]	813,484 (67.6%) [a]
Number (%) of resident episodes in the numerator	56,786 (4.7%) [a]	56,013 (4.7%) [a]
Number (%) of facilities able to report this QM	11,352 (72.4%) [b]	11,283 (71.9%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay sample (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\ )

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.7-6 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. For each pair of quarters, facility rankings tended to be somewhat less stable than for other quality measures. From quarter to quarter, just over half of facilities maintained their ranking, whereas a substantial portion (at least 16.1 percent) saw their percentile rank change by more than 30 percent. These ranking shifts are likely due to this being a low-

prevalence measure: the changes in actual scores are relatively small, as shown in the next analysis.

Table 4.7-7 presents the results from an analysis of the change in facility scores for this quality measure between the third and fourth quarters of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 4.7 percent). Overall, the average change in scores for this measure was a decline of 0.1 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (72.5 percent saw changes of less than one standard deviation), whereas fewer than 1 percent saw changes of more than three standard deviations.

Table 4.7-6
Distribution of change in facility rank from one quarter to the next, QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	5,384 (50.8%)	5,466 (50.7%)	5,508 (50.4%)
Between 1 and 2 deciles	2,178 (20.5%)	2,143 (19.9%)	2,147 (19.6%)
Between 2 and 3 deciles	1,338 (12.6%)	1,350 (12.5%)	1,404 (12.8%)
More than 3 deciles	1,709 (16.1%)	1,831 (17.0%)	1,880 (17.2%)
Total	10,609	10,790	10,939

NOTES:

Total number of facilities in each column reflects all the facilities that could report the measure in both quarters.

Percentages may not total 100.0 because of rounding.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 to Quarter 4, 2011 (\\Rtimas04\hser\Project\0211942\001 MIDS-NHQ\004 Testing and Validation\common\ykaganova\db\quarter_4_5\db155_request\db155_request.log)

Table 4.7-7
Change in facility score from one quarter to the next, QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
10,939	-0.1%	4.0%	0.5%	2.4%	10.8%	34.7%	37.8%	11.2%	2.3%	0.4%

NOTES:

Number of facilities reporting (10,939) reflects facilities that reported this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures, should be correlated). Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures where applicable. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) and the facility's percentile rank on QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) in Quarter 4, 2011, given that both are concerned with the outcome of pressure ulcers. We found that the correlation between the nursing home's percentile rank on QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) and percentile rank on QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) was small (correlation coefficient = .148) and statistically significant ($p < .001$). [20] The weak correlation may be due to the different specifications for the short-stay and long-stay QMs, i.e., the short-stay QM is an incidence measure whereas the long-stay QM is a prevalence measure. It may also suggest that the care processes for preventing pressure ulcers and/or the patient characteristics in the short-stay versus long-stay populations are very different, and that facilities may have more specialized expertise in dealing with only one of the patient groups.

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether State characteristics might be a source of facility score variation for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) and thus a potential threat to the measure's validity, we conducted a one-way analysis of variance (ANOVA). We also examined the interquartile difference between the mean State-level scores across States. The proportion of variance in this measure explained by the State in which facilities are located is 6.5 percent and significant [$F(52, 11{,}301) = 15.6$, $p < .001$].[21] The difference between the mean State-level scores for States at the 25th percentile and the 75th percentile is 2.1 percentage points. Thus, although the majority of the variance in QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) scores is due to factors other than geography, a small and significant proportion of the variance is explained by

[20] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

[21] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

the respective States in which the nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns of missing data indicate that certain types of residents tend to have assessments with missing data in ways that affect the calculation of a quality measure, then that measure may not be capturing processes and outcomes of care for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

Missing data for this QM were not prevalent and therefore unlikely to present a threat to the validity of this QM. After excluding admission, PPS 5-day, and readmission assessments (these were 18,102 resident episodes, 1.5 percent of the long-stay population), there were 387,218 residents (32.3 percent of the long-stay population) excluded because they did not meet the high-risk criteria. Of the remaining residents, only 699 (less than one-tenth of 1 percent) were excluded for missing data.

Table 4.7-8 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure.

We further examined the possible relationship between missing data and QM scores. ***Table 4.7-9*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure; they are near-zero across quartiles; however, there is a significant but very weak correlation between missing data rates and QM scores ($r = 0.047$, $p < 0.001$). Missing data account for an extremely small proportion (approximately 0.2 percent) of the variance, thus, for this measure, missing data does not have a substantial effect on scores. There are almost no missing data on items used to calculate this measure (the facility mean missing rate is 0.1 percent, and at least 90 percent of facilities have no missing data), indicating that having missing data is not a threat to validity.

Table 4.7-8
Distribution of facility-level missing rate for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,307	0.1%	1.2%	0.0%	0.0%	0.0%	0.0%	0.0%

NOTES:

n = number of facilities that have data for numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request.log)

Table 4.7-9
Distribution of facility-level missing rate for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) by quartile by QM score

Quartile by QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,868	0.1%	1.0%	0%	0%	0%	0%	0%
26%–50%	3,777	0.1%	0.8%	0%	0%	0%	0%	0%
51%–75%	3,806	0.1%	0.9%	0%	0%	0%	0%	0%
76%–100%	3,803	0.1%	1.9%	0%	0%	0%	0%	0%

NOTES:

Total *n* = number of facilities that have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2) the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.7-10.*** The mean nursing home–level QM score and distributions changed slightly before and after excluding discharge assessments from the set of eligible target assessments. The average facility score decreased by 0.1 percent (with a median change of 0 percent) when discharge assessments were excluded. Thus, the use of the discharge assessment in the QM score calculations does not meaningfully alter the distribution of scores among facilities.

Table 4.7-10
Distribution of facility-level score change after excluding discharge assessment for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
11,283	-0.1%	1.1%	-1.3%	0.0%	0.0%	0.1%	0.9%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

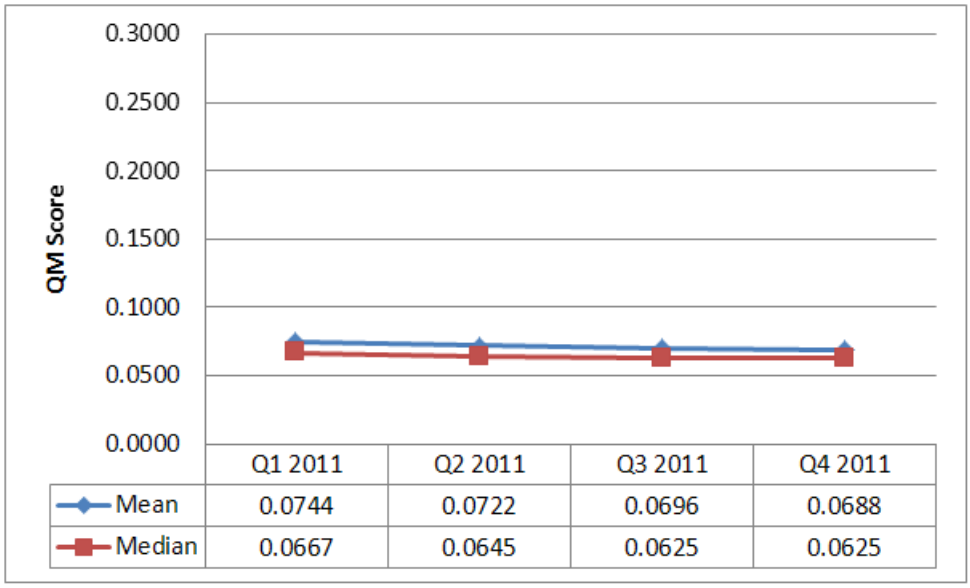
SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request_v1\db135_request_v1.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address this interest in whether seasonal variation might play a role in the score for QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay), we examine the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.7-1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, was 7.4 percent, 7.2 percent, 7.0 percent, and 6.9 percent. The quarterly national median was 6.7 percent, 6.5 percent, 6.3 percent, and 6.3 percent. Both the mean fell slowly over the four quarters, and the median fell slowly over the first 3 quarters, and remained constant from Quarter 3, 2011, to Quarter 4, 2011. With only 12 months of data available, it is too early to ascertain whether the changes in national QM score over time are consistent with seasonal variation; this will need to be confirmed when multiple years of data become available.

Figure 4.7-1
Seasonal (quarterly) variation in QM #0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay)



	Q1 2011	Q2 2011	Q3 2011	Q4 2011
Mean	0.0744	0.0722	0.0696	0.0688
Median	0.0667	0.0645	0.0625	0.0625

Analysis date:

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 through Quarter 4, 2011 (\qm_quarter_1_2\complete\nh_013_10.log f \qm_quarter_2_3\ complete\nh_013_10.log \qm_quarter_3_4\ complete\nh_013_10.log \qm_quarter_4_5\ complete\nh_013_10.log)

4.7.5 Risk Adjustment

This quality measure is risk adjusted using denominator exclusions. Residents at low risk for pressure ulcers are excluded from the denominator. High risk is defined as any of the

following conditions: impaired bed mobility/transfer (G0110A/B); comatose (B0100); or malnutrition/at risk for malnutrition (I5600). As previously presented in Table 4.7-1, about two-thirds of the long-stay residents meet the high-risk criteria and are therefore included in the denominator. It was suggested in the public comment period that the risk adjusters used for the short-stay pressure ulcer measure should be considered for model-based risk adjustment (diabetes, peripheral vascular disease, and bowel incontinence). However, based on the QM in Quarter 3, 2011, the preliminary risk-adjustment model with the four risk adjusters does not show sufficient predictive power in the long-stay sample, suggested by a C-statistic of 0.575 (Pseudo R-square = 0.010).[22] Further analyses should investigate other combinations of risk adjusters, either through exclusions or including covariate adjustments in the QM model. Additional resident characteristics to be considered include hospice use, history of pressure ulcer, hip fracture, spinal cord injury, and other mobility-limiting diagnoses.

4.8 QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)

4.8.1 Summary of Findings

QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) reports the percentage of long-stay residents who either received the influenza vaccine during the current or most recent influenza season (either in the facility or outside the facilities) or offered and declined the vaccine or were ineligible due to contraindication(s). This is an important measure of quality of care in the nursing facility, given that morbidity and mortality related to influenza are often reported in conjunction with data regarding pneumonia, and together frequently lead to death in the elderly population.

This subsection presents descriptive analyses and test results for this QM. We found that 98.5 percent of the long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 88 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 89.1 percent. The QM score varies across facility with a standard deviation of 13.7 percent and interquartile range of 14.8 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

As reliability checks of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. For about 44 percent of the facilities, the QM score increased by less than one standard deviation. For 38.5 percent of facilities, the QM score decreased by less than one standard deviation. From Quarter 3 to Quarter 4, 2011, about half of facilities remained within the same decile, but 17.6 percent of facilities had rank changes of more than three deciles.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a care process group. The hypothesis was that changes in one QM within a measure group (e.g., vaccination QMs) should be correlated with changes in other measures

[22] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 3, 2011 (\quarter_3_4_complete\ntz10_request.log)

because they reflect similar care processes. Findings from this analysis showed moderate to high correlations among the vaccination QMs. We also analyzed the potential geographic (State) and seasonal (quarter) variations in this QM. The State in which the facility is located explains about 3.3 percent of the variation in this QM. The QM score slightly declined from Quarter 1 to Quarter 4 in 2011.

4.8.2 Background and Introduction to QM

The quality measure reports the percentage of long-stay residents who were appropriately given the influenza vaccination during the current or most recent influenza season. The measure is reported as the aggregate of three separately calculated submeasures to reflect the various ways that a resident may be "appropriately" given the influenza vaccination during the current or most recent influenza season. The three submeasures are also calculated and reported separately to harmonize with the NQF and CDC reporting structures. The three submeasures are

- resident received the influenza vaccine during the current or most recent influenza season, either in the facility or outside the facility;
- resident was offered and declined the influenza vaccine; and
- resident was ineligible to receive the influenza vaccine due to contraindication(s) (e.g., anaphylactic hypersensitivity to eggs or other components of the vaccine, history of Guillain-Barre Syndrome within 6 weeks after a previous influenza vaccination, bone marrow transplant within the past 6 months).

A. Background for This QM

Morbidity and mortality data related to influenza are often reported in conjunction with data regarding pneumonia. According to the CDC, pneumonia and influenza was the seventh most common cause of death for people aged 65 and older in the United States (Centers for Disease Control and Prevention, 2009). In 2009, influenza and pneumonia combined caused 43,465 deaths in people over the age of 65, with 638 deaths caused from influenza alone. In addition to being at risk for primary illness, frail elderly are especially vulnerable and subject to complications of influenza. In 2004, there were approximately 123,000 deaths with influenza and pneumonia mentioned on the death certificate as a secondary cause of death (Gorina et al., 2008).

According to the CDC, more than 200,000 people are hospitalized in the United States each year as a result of complications from influenza (Centers for Disease Control and Prevention, 2008). The average hospital stay was approximately 5.3 days at a cost of $6,900 per stay (Milenkovic et al., 2006). Further, the death rate from influenza among the elderly ranges from 1.1 to 3.6 per 100,000, with risk increasing with age (Centers for Disease Control and Prevention, 2009). The death rate for influenza and pneumonia in people 65 to 74 years old is three times that of a person 55 to 64; and for a person over the age of 85, the death rate is 35 times that (Centers for Disease Control and Prevention, 2009). Among adults aged 65 and older, approximately 67.4 percent were vaccinated during 2010, which is below the Healthy People 2020 target of 90 percent for this age group (Centers for Disease Control and Prevention; U.S. Department of Health and Human Services, 2000). For adults in nursing facilities, in 2004, the National Nursing Home Survey found that 62.9 percent of residents were vaccinated against

influenza (National Nursing Home Survey, 2004). The most recent Nursing Home Compare data indicate that the number has risen, to 85 percent for short-stay residents and 92 percent for long-stay residents (Centers for Medicare & Medicaid Services, 2012).

This measure is intended to encourage nursing facilities to focus on this important aspect of clinical care by assessing residents on the status of their seasonal flu vaccine immunization and to provide immunization as appropriate.

There are two QMs for immunization: one for the long-stay and one for the short-stay populations. This is because they are two distinct populations with inherent clinical differences that are easily captured by these two length-of-stay categories. Furthermore, nursing facilities that serve a short-stay population have a limited time frame to assess and provide vaccinations compared with the long-stay population.

B. Summary of Differences between MDS 2.0 and MDS 3.0 QM Definitions

QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long-Stay) reports the percentage of all long-stay residents who were assessed and appropriately given the seasonal influenza vaccine during the influenza season as reported on the target MDS assessment: (1) received the influenza vaccine during the most recent influenza season, either in the facility (O0250A = 1) or outside the facility (O0250C = 2; or (2) were offered and declined the influenza vaccine (O0250C = 4); or (3) were ineligible due to contraindications(s) (O0250C = 3). The numerator includes residents with a target assessment (OBRA MDS 3.0 assessment [A0310A = 01, 02, 03, 04, 05, 06], or PPS assessment [A0310B = 01, 02, 03, 04, 05, 06] or discharge assessment [A0310.F = 10, 11]) during the influenza reporting period. When a vaccination is completed at the end of the influenza season, the next opportunity to report the vaccination may be after the season is over. Extending the seasonal influenza vaccination reporting period through 12 months allows for the capture of those late-season vaccinations reported after the season ends. Note that residents are not excluded from the measure sample if there is a missing response on the item indicating whether the influenza vaccine was received (O0250A), rather these residents are assumed to have not received the vaccine and counted in the denominator. Additionally, for residents who did not receive the vaccination (O0250A = 0), if the item indicating the reason it was not administered was left missing (O0250C), it is assumed that there was no valid reason for the resident not to receive the vaccine and the resident is counted in the denominator of the measure.

The underlying MDS 3.0 items used to construct this measure did not change from the MDS 2.0 measure. Minor item changes to clarify the item included the addition of a “none of the above” category for the reason the vaccine was not given (O0250C=9). In addition, the MDS 3.0 measure now includes individuals who refuse to be vaccinated or have medical contraindications to vaccination, potentially increasing the number of residents who might be counted in the aggregated QM numerator and denominator compared with the MDS 2.0 QM.

Note also that the overall sample definitions changed from MDS 2.0 to MDS 3.0. In the MDS 2.0 measure, residents were included in the chronic care sample if they had a full or quarterly MDS in the target quarters; the measure was calculated based on just these two types of assessments. In the MDS 3.0 measure, the analogous sample is for the long-stay residents,

defined as residents with more than 100 cumulative days in a facility, and whose assessments may be discharge, 5-, 14-, 30-, 60-, 90-day PPS assessments; or admission, quarterly, annual, or significant change OBRA assessments.

C. Summary of Analyses

The following sections summarize the basic analyses performed for this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- number of assessments that are included and excluded from the numerator of the QM based on sample restrictions;
- whole sample prevalence of items that are related to the QM; and
- findings regarding variability, reportability, reliability, and validity.

4.8.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 1,203,663 assessments included in the denominator (approximately 100 percent of all assessments). A total of 17,965 resident episodes (1.5 percent of the total number of long-stay resident episodes) were excluded from the denominator of QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay). ***Table 4.8-1*** shows the proportion of resident episodes excluded from the denominator for this measure's exclusion criterion. For this measure, residents are excluded only if they were not in the facility during the current or most recent influenza season. A total of 63 (0.0 percent) assessments met this criterion.

Table 4.8-1
Target assessments used in the calculation of QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	1,185,761	98.5%
Excluded—Exclusion 1	17,965	1.5%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Resident was not in facility during the current or most recent influenza season (O0250C = [1]).

[a] Column values may not add up to total because a resident episode can meet more than one exclusion criterion. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

As illustrated in ***Table 4.8-2,*** of the total target assessments not excluded from the denominator, more than 71 percent indicated that the resident had received the influenza vaccine during the most recent influenza season (O0250A = 1). Of those residents who did not receive the vaccine in the facility, almost 4 percent of residents received the influenza vaccine during the most recent influenza season outside the facility (O0250C = 2); almost 12 percent of residents were offered and declined the influenza vaccine (O0250C = 4); and 1 percent of residents were ineligible to receive the vaccine owing to contraindication(s) (O0250C = 3). Other response categories to O0250C are missing [—]; not offered [5]; inability to obtain vaccine due to declared shortage [6]; or none of the above [9].

Table 4.8-2
Frequency and percentage responses on items used to define the numerator for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay)

Item	Category	Frequency	Percentage	Cumulative %
O0250A	— (missing)	11,762	1.0%	1.0%
O0250A	0 = No	335,753	27.9%	28.9%
O0250A	1 = Yes	856,211	71.1%	100.0%
O0250A	Total	1,203,726	—	—
O0250C	— (missing)	16,366	1.4%	1.4%
O0250C	1 = Resident not in facility during this year's flu season	17,965	1.5%	2.9%
O0250C	2 = Received outside of this facility	42,703	3.6%	6.4%
O0250C	3 = Not eligible – medical contraindication	12,104	1.0%	7.4%
O0250C	4 – Offered and declined	142,828	11.9%	19.3%
O0250C	5 – Not offered	61,004	5.1%	24.3%
O0250C	6 – Inability to obtain vaccine due to a declared storage	1,282	0.1%	24.5%
O0250C	9 – None of the above	53,263	4.4%	28.9%
O0250C	^ (skipped)	856,211	71.1%	100.0%
O0250C	Total	1,203,726	—	—

NOTE: O0250a = Resident received vaccine in-facility for the current influenza season); 0250c = If influenza vaccine not received, state reason.

Analysis date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_015_10.log)

4.8.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with perfect scores (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes), which may indicate a "ceiling effect." ***Table 4.8-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay) is 89.1 percent with a standard deviation of 13.7 percent. This QM has an interquartile range of 14.8 percent. About 21.2 percent of facilities have scores of 100 percent (perfect scores). The results for variability for the three submeasures are presented in ***Table 4.8-3a*** to ***4.8-3c***.

Table 4.8-3
QM score distribution for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
13,795	89.1%	13.7%	70.6%	84.0%	94.4%	98.8%	100%	21.2%	14.8%

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting of this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_015_10.log)

Table 4.8-3a
QM score distribution for submeasure QM #0681A Percent of Residents Who Received the Seasonal Influenza Vaccine (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
13,795	76.2%	16.0%	54.0%	67.8%	80.0%	88.1%	93.2%	20.3%

Analysis date: 7/3/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_015.log)

Table 4.8-3b
QM score distribution for submeasure QM #0681B Percent of Residents Who Were Offered and Declined the Seasonal Influenza Vaccine (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
13,795	11.8%	9.0%	2.9%	5.7%	9.8%	15.6%	23.1%	10.0%

Analysis date: 7/3/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting of this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_015.log)

Table 4.8-3c
QM score distribution for submeasure QM #0681C Percent of Residents Who Did Not Receive, Due to Medical Contraindication, the Seasonal Influenza Vaccine (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
13,795	1.0%	1.6%	0.0%	0.0%	0.0%	1.7%	2.9%	1.7%

NOTES:

n = number of facilities that meet minimum requirements for public reporting of this quality. measure.

QM scores are reported at the facility level.

Analysis date: 7/3/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_015.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 long-stay residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examined the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.8-4*** (second column) shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 88.0 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay) using MDS 3.0. This is a slight decline from the MDS 2.0 measure, which 89.5 percent of facilities were able to report in Quarter 1, 2006 (Brega et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more

complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and that 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may affect the reportability of short-stay QMs to a greater degree than the long-stay QMs.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay), we examined changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.8-4.*** After excluding discharge assessments from the target assessment sample, reportability was largely unchanged for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay) (88.0 percent with discharge assessments vs. 87.8 percent without discharge assessments).

Table 4.8-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator	1,185,761 (98.5%) [a]	1,180,813 (98.1%) [a]
Number (%) of resident episodes in the numerator	1,053,846 (87.5%) [a]	1,046,316 (86.9%) [a]
Number (%) of facilities able to report this QM	13,795 (87.9%) [b]	13,779 (87.8%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay sample (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.8-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. Between Quarters 1 and 2 of 2011, facility rank was stable, with 79.1 percent of

facilities maintaining the same rank (within 10 percentile points) and just 4.1 percent showing moves of more than three deciles. For the other two transitions, however, facility rank was less stable, with far fewer (44.8 percent and 52.2 percent, respectively) facilities maintaining their rank. However, as we see with subsequent analyses (below), these changes in rank were brought about by very small changes in scores, indicating that QM scores for this measure are clustered closely together.

Table 4.8-5
Distribution of change in facility rank from one quarter to the next, QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	10,743 (79.1%)	6,061 (44.8%)	7,047 (52.2%)
Between 1 and 2 deciles	1,570 (11.6%)	5,400 (39.9%)	2,411 (17.8%)
Between 2 and 3 deciles	718 (5.3%)	1,043 (7.7%)	1,681 (12.4%)
More than 3 deciles	552 (4.1%)	1,033 (7.6%)	2,375 (17.6%)
Total	13,583	13,537	13,514

NOTES:

Total number of facilities in each column reflects all the facilities that could report the measure in both quarters.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.8-6 presents the results from an analysis of the change in facility scores for this quality measure between the third and the fourth quarters of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 18.4 percent). Overall, the average change in scores for this measure was a decline of 0.3 percentage point. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (82.0 percent saw changes of less than one standard deviation), whereas just under 3 percent saw changes of more than three standard deviations.

Table 4.8-6
Change in facility score from one quarter to the next, QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations:	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
13,514	-0.3%	18.4%	0.3%	1.5%	7.8%	38.5%	43.5%	4.3%	1.5%	2.6%

NOTES:

Number of facilities reporting (13,514) reflects facilities that reported this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures, should be correlated). Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures where applicable. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay) and the facility's percentile rank on QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay), QM #0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay), and QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) in Quarter 4, 2011, because all of these QMs measure the vaccination process. We found that the correlation between the nursing home's percentile rank on QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) and percentile rank on QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) was moderate (correlation coefficient = .557) and statistically significant ($p < 0.001$). The correlations between this measure and the short-stay and long-stay pneumococcal vaccination measures—QM #0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay) and QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)—were low (0.342 and 0.392, respectively) but still statistically significant ($p < .001$ in both cases).[23] The moderate correlation possibly indicates that some facilities have systems in place for identifying patients in need of updating their vaccinations and providing them, whereas other facilities provide vaccinations on an as-needed or upon-request basis without a formal system in place. It may also suggest that facilities' attention to vaccines is vaccine type specific (influenza or pneumococcal).

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether State characteristics might be a source of facility score variation for QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) and thus a potential threat to the measure's validity, we conducted a one-way analysis of variance (ANOVA). We examined the interquartile difference between the mean State-level scores across States. The proportion of variance in this measure explained by the State in which facilities are located is 3.3 percent and

[23] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

significant [$F(51, 13{,}743) = 9.2$, $p < .001$].[24] The difference between the mean State-level scores for States at the 25th percentile and the 75th percentile is 4.2 percentage points. Thus, although the majority of the variance in QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) scores is due to factors other than geography, a small and significant proportion of the variance is explained by the respective States in which nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns indicate that certain types of residents tend to have assessments with missing data in ways that affect the calculation of a quality measure, then that measure may not be capturing processes and outcomes for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

In Quarter 4 of 2011, there were 17,965 residents (1.5 percent of the long-stay population) who were excluded from the construction of this measure because they were not in the facility during the then-current or most recent influenza season (indicated by O0250C = [1]). For this measure, values of [0] are imputed for missing data. Thus, no long-stay resident was excluded from the QM calculation due to missing data. ***Table 4.8-7*** shows the distribution of facility-level missing-data rates for items used to construct this measure. Missing data were rare for this measure, with a mean facility-level missing rate of 1.8 percent. This distribution of these rates also has a median of 0 percent and a 90th percentile value of 4.9 percent, suggesting that relatively few facilities may be driving the national average.

Table 4.8-7
Distribution of facility-level missing rate for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,366	1.8%	4.9%	0%	0%	0%	2.0%	4.9%

NOTES:

n = number of facilities that have data for numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

We further examined the possible relationship between missing data and QM scores. ***Table 4.8-8*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure: they are essentially identical across quartiles. There is an extraordinarily weak but significant correlation between missing-data rates and QM scores ($r = -0.050$, $p < 0.001$). As stated above, values of 0 are imputed for missing data on this measure. Given that there is a small amount of missing data, it makes sense that the imputation decreases scores by a small amount. However, missing data accounts for far too little of the variance (0.25 percent) to have a substantial impact on QM scores for this measure.

[24] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011(\quarter_4_5\db160_request)

Table 4.8-8
Distribution of facility-level missing rate for QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay) by quartile of QM score

Quartile of QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,842	2.1%	5.8%	0.0%	0.0%	0.0%	2.1%	5.1%
26%–50%	3,841	1.7%	3.5%	0.0%	0.0%	0.0%	2.0%	5.0%
51%–75%	3,845	1.6%	3.1%	0.0%	0.0%	0.0%	2.0%	4.4%
76%–100%	3,838	2.0%	6.3%	0.0%	0.0%	0.0%	1.7%	5.3%

NOTES:

Total *n* = number of facilities that have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2) the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.8-9.*** The mean nursing home–level QM score and distributions change slightly before and after excluding discharge assessments from the set of eligible target assessments. The average facility score decreased by 0.3 percent (with a median change of 0 percent) when discharge assessments are excluded. Thus, the use of the discharge assessment in the QM score calculations does not meaningfully alter the distribution of scores among facilities for this measure.

Table 4.8-9
Distribution of facility-level score change after excluding discharge assessment for QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
13,779	-0.3%	1.7%	-1.0% -2.0%	0% -0.4%	0.0%	0.0%	1.0%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

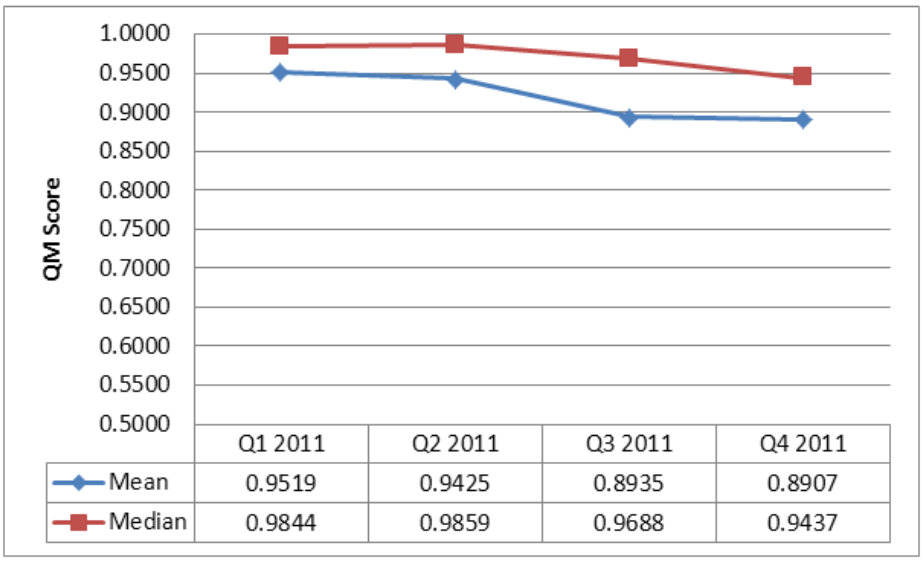
SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request_v1\db135_request_v1.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address this interest in whether seasonal variation might play a role in the score for QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay), we examined the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.8-1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011 and Quarter 4, 2011, was 95.2 percent, 94.3 percent, 89.4 percent, and 89.1 percent. The quarterly national median was 98.4 percent, 98.6 percent, 96.9 percent, and 94.4 percent. Although both mean and median were quite high for all quarters, the mean decreased over each quarter, whereas the median increased from Quarter 1, 2011, to Quarter 2, 2011, and decreased for subsequent quarters. Compared with other QMs, this QM appears more subject to seasonal variation. This is, however, as expected owing to flu season. With only 12 months of data available, it is too early to ascertain whether the changes in national QM score over time are consistent with seasonal variation; this will need to be confirmed when multiple years of data become available.

Figure 4.8-1
Seasonal (quarterly) variation in QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay)



	Q1 2011	Q2 2011	Q3 2011	Q4 2011
Mean	0.9519	0.9425	0.8935	0.8907
Median	0.9844	0.9859	0.9688	0.9437

Analysis date:

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 - Quarter 4, 2011 \qm_quarter_1_2\complete\nh_015_10.log f \qm_quarter_2_3\ complete\nh_015_10.log \qm_quarter_3_4\ complete\nh_015_10.log \qm_quarter_4_5\ complete\nh_015_10.log

4.8.5 Risk Adjustment

QM #0681 Percent of Residents Assessed and Appropriately Given the Influenza Vaccination (Long Stay) was endorsed by NQF without denominator exclusion and model-based risk adjustment.

4.9 QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)

4.9.1 Summary of Findings

QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay) reports the percentage of long-stay residents whose pneumococcal vaccination is up-to-date or who were offered and declined the vaccine or were ineligible due to contraindication(s). Because this QM does not have denominator exclusions, all short-stay residents are included in the denominator. About 88 percent have 30 or more long-stay resident episodes included in the denominator and are able to report this QM. Among facilities able to report, the mean facility QM score was 93.8 percent. The QM score varied across facility, with a standard deviation of 12.4 percent and interquartile range of 6.4 percent. About 40 percent of facilities have perfect scores (i.e., 100 percent). These findings indicate that most facilities perform well in this aspect of care. The relatively large standard deviation suggests that this QM can be particularly useful to identify facilities with poor quality.

To determine the reliability of this QM, we examined the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The large majority of facility score changes, either improvements or declines, were within one standard deviation. A very low proportion of changes were greater than three standard deviations. A similar pattern was found for facility rank changes. More than 80 percent of facilities remained within the same decile from quarter to quarter. Only about 2 percent of facilities shift more than three deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a care process group. The hypothesis was that changes in one QM within a measure group (e.g., vaccination QMs) should be correlated with changes in other measures because they reflect similar care processes. Findings from this analysis showed that this QM is highly correlated with the short-stay pneumococcal vaccination QM (correlation $r = 0.690$) and the short-stay influenza vaccination QM (correlation $r = 0.533$). This QM is significantly but moderately correlated with the long-stay influenza vaccination QM (correlation $r = 0.392$). We also analyzed the potential geographic (State) and seasonal (quarter) variations in this QM. The State in which the facility is located explains about 4.5 percent of the variation in this QM. The QM score was largely stable from quarter to quarter in 2011.

4.9.2 Background and Introduction to QM

QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay) is based on data from MDS 3.0 assessments of long-stay nursing facility residents. The measure reports the percentage of all long-stay residents who were assessed and appropriately given the pneumococcal vaccination (PPV) as reported on the target

MDS assessment (OBRA, PPS, or discharge) during the 12-month reporting period. This measure is harmonized with NQF's quality measure on pneumococcal immunizations (National Quality Forum, 2008). The MDS 3.0 definitions have been changed to conform to the NQF standard. The NQF used current guidelines from the Advisory Committee on Immunization Practices (ACIP) and others to guide decisions on all parameters for the harmonized measures (ACIP, 1997). The NQF standard specifications were harmonized to achieve a uniform approach to measurement across settings and populations, addressing who is included or excluded in the target denominator population, who is included in the numerator population, and time windows for measurement and vaccinations. Long-stay residents are those residents who have been in the nursing home facility for at least 100 days. The measure is restricted to the population with long-term care needs and does not include the short-stay population who are discharged within 100 days of admission.

This measure is intended to encourage nursing facilities to focus on this important aspect of clinical care by assessing residents on the status of their pneumococcal vaccine immunization and to provide immunization as appropriate.

A. Background for This QM

This is a very important measure of quality of care in the nursing facility. Morbidity and mortality related to pneumonia are often reported in conjunction with data regarding influenza, and together frequently lead to death in the elderly population. According to the Centers for Disease Control and Prevention (CDC), pneumonia and influenza was the seventh most common cause of death for persons aged 65 and older in the United States (Centers for Disease Control and Prevention, 2009). In 2009, pneumonia caused 50,774 deaths, 85 percent of which were in people over the age of 65 (Centers for Disease Control and Prevention, 2009). In addition to being a primary cause of death, pneumonia and influenza are often complications of other illnesses. In 2004, there were approximately 123,000 deaths with influenza and pneumonia mentioned on the death certificate as a secondary cause of death (Gorina et al., 2008).The death rate for influenza and pneumonia in people 65-74 years old is three times that of a person 55-64, and for a person over the age of 85, the death rate is 35 times that (Centers for Disease Control and Prevention, 2009). According to CDC, pneumococcal disease kills more people in the United States each year than all other vaccine-preventable diseases combined (Centers for Disease Control and Prevention, 1997).

Older people and persons with chronic health conditions are at high risk for pneumococcal disease. However, there is a demonstrated gap in performance in vaccination among adults aged 65 years and older. In 2010, 68.6 percent of adults over the age of 65 reported having a pneumococcal vaccine (Centers for Disease Control and Prevention, 2010), less than the 90 percent goal set by Healthy People 2020 (U.S. Department of Health and Human Services, 2012). Vaccinations of nursing facility residents can prevent or lower the risk of residents becoming seriously ill. According to data from Nursing Home Compare, from 2007 to 2012, 91 percent of short-stay nursing home residents and 85 percent of long-stay residents have received the pneumococcal vaccine, up from 45.4 percent in 2004 (Centers for Medicare & Medicaid Services) (National Nursing Home Survey, 2004).

Healthy People 2020 includes Objective IID-13, for institutionalized adults, of a 90 percent vaccination rate for pneumonia in 2020 (U.S. Department of Health and Human

Services, 2012). Hospitalization rates for pneumonia-related stays for the elderly population have been increasing over the past 15 years, and among those 85 and older, at least 1 in 20 elderly persons were hospitalized each year because of pneumonia (Fry et al., 2005). In 2005, Medicare paid an average of $6,342 per hospital discharge for pneumonia-related short-stay hospitalizations; the average length of stay was 6.1 days. In 2007, There were 610,000 hospital discharges for pneumonia in patients over the age of 65, with an average length of stay of 5.0 days (Centers for Medicare & Medicaid Services, 2007).

B. Summary of Differences between MDS 2.0 and MDS 3.0 QM Definitions

QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) reports the percentage of all long-stay residents who were assessed and appropriately given the seasonal pneumococcal vaccine during the influenza season. Residents are counted in the numerator 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. Long-stay residents are included in the numerator if they meet any of the following criteria on the target MDS 3.0 assessment (A0310.A=01,02,03,04,05,06; A0310B = 01,02,03,04,05,06; or A0310.F= 10,11) during the 12-month reporting, and: (1) have an up-to-date PPV status (item O0300.A= 1); or (2) were offered and declined the vaccine (item O0300.B=2); or (3) were ineligible due to medical contraindication(s) (i.e., anaphylactic hypersensitivity to components of the vaccine, bone marrow transplant within the past 12 months, or receiving a course of chemotherapy within the past 2 weeks) (item O0300B=1). Note that residents are not excluded from the measure sample if there is a missing response on the item indicating whether their PPV status is up-to-date (O0300A), rather it is assumed that these residents do not have an up-to-date PPV vaccination and they are counted in the denominator. Additionally, for residents who do not have an up-to-date PPV vaccination (O0300A = 0), if the item indicating the reason it was not administered was left missing (O0300B), it is assumed that there was no valid reason for the resident not to receive the vaccine and the resident is counted in the denominator of the measure.

There are no resident level exclusions from the denominator. The denominator includes all long-stay residents who meet the following criteria: (1) the target MDS 3.0 assessment is an OBRA assessment (item A0310A = 01, 02, 03, 04, 05, 06) or a PPS assessment (A0310B = 01, 02, 03, 04, 05, 06.) or a discharge assessment (item A0310F = 10,11) with discharge date (item A2000) during the 12-month target period.

This is a change from the MDS 2.0 nursing home quality measure. The underlying MDS items used to construct this measure did not change from the MDS 2.0 to the MDS 3.0. The changes made to the MDS 3.0 regarding the vaccine items were relatively minor wording changes meant to clarify the item. The reasons for the pneumococcal vaccine not being received are identical to the MDS 2.0 item. Finally, as with the MDS 2.0 pneumococcal vaccination measures, the MDS 3.0 calculates separate measures for the short and long-stay populations. There are inherent differences in nursing facility's being responsible for assessing and/or providing vaccines for these distinct populations. For the short-stay population, nursing facilities have less time to assess and/or provide the vaccine than for the long-stay population. As a result, nursing facilities' vaccination rates for post-acute care populations should not be compared to

rates for long-term care populations. Separating them recognizes these differences in vaccination rates.

This measure is based on the NQF's National Voluntary Standards for Influenza and Pneumococcal Immunizations, which include resident refusal and ineligibility in the numerator and denominator. The main difference between the MDS 2.0 and MDS 3.0 measures is that the overall sample definitions also changed. In the MDS 2.0 QM, residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments. In the MDS 3.0 QM, the analogous sample is the long-stay residents, defined as a resident with more than 100 cumulative days in facility; the associated assessments may be any of the following (unless specifically excluded as described above): discharge; 14-, 30-, 60-, or 90-day PPS assessments; or quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections summarize the basic analyses performed for this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- number of assessments that are included and excluded from the numerator of the QM based on sample restrictions;
- whole sample prevalence of items that are related to the QM; and
- findings regarding variability, reportability, reliability, and validity.

4.9.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

There were 1,203,726 episodes included in the numerator (100 percent)—there are no exclusions in the calculation of QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay). Missing numerator values are imputed as "0." Thus, 100 percent of the 1,203,726 target assessments in the long-stay sample were included in the denominator for this measure (***Table 4.9-1***).

Table 4.9-1
Target assessments used in the calculation of QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	1,203,726	100%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

There are no exclusions for this measure.

[a] Column values may not add up to total because a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

Table 4.9-2 describes the responses to the key MDS 3.0 item used to calculate this QM. Item O0300a asks whether the resident's pneumococcal vaccination is up-to-date. Nearly 79 percent of the sample shows that the vaccination is up-to-date (O0300a=1) and would be counted in the numerator based on this information. More than 14 percent of residents were offered the vaccine and declined (O300b=2), which would also count them in the numerator.

Table 4.9-2
Responses to O0300a (Pneumonia Vaccine Up-to-Date) and O0300b (Reason for Not Receiving Pneumonia Vaccine) on target assessments

Item	Entry	Frequency	Percentage	Cumulative %
O0300a	- = missing	21,579	1.8%	1.8%
O0300a	0 = Not up-to-date	236,153	19.6%	21.4%
O0300a	1 = Up-to-date	945,994	78.6%	100.0%
O0300a	Total	1,203,726	—	—
O0300b	- = missing	33,284	2.8%	2.8%
O0300b	1 = Not eligible – medical contraindication	9,597	0.8%	3.6%
O0300b	2 = Offered and declined	170,242	14.1%	17.7%
O0300b	3 = Not offered	44,609	3.7%	21.4%
O0300b	^ = Skipped because vaccine is up-to-date	945,994	78.6%	100.0%
O0300b	Total	1,203,726	—	—

NOTE: O0300a = Pneumococcal vaccination up to date; O0300b = If Pneumococcal vaccine not received, state reason.

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\ \nh_017_10.log)

4.9.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with perfect scores (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes) which may indicate a "ceiling effect." ***Table 4.9-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) is 93.8 percent with a standard deviation of 12.4 percent. This QM has an interquartile range of 6.4 percent. About 39.5 percent of facilities have scores of 100 percent (perfect scores). The results for variability for the three sub-measures are presented in ***Table 4.9-3a*** to ***4.9.3c***.

Table 4.9-3
QM score distribution for QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
13,870	93.8%	12.4%	82.5%	93.6%	98.4%	100%	100%	39.4%	6.4%

Analysis date: 5/29/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting of this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5 \nh_017_10.log)

Table 4.9-3a
QM score distribution for submeasure QM #0683A Percent of Residents Who Received the Pneumococcal Vaccine (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
13,870	79.1%	18.0%	55.1%	71.9%	84.3%	91.7%	96.0%	19.8%	13,870

Analysis date: 7/3/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting of this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_017.log)

Table 4.9-3b
QM score distribution for submeasure QM #0683B Percent of Residents Who Were Offered and Declined the Pneumococcal Vaccine (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
13,870	13.9%	13.1%	2.2%	5.2%	10.3%	18.4%	30.0%	13.2%	13,870

Analysis date: 7/3/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting of this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db178_request\db178_request_017.log)

Table 4.9-3c
QM score distribution for submeasure QM #0683C Percent of Residents Who Did Not Receive, Due to Medical Contraindication, the Pneumococcal Vaccine (Long Stay)

n	Mean score	Std Dev.	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile	% of Facilities with "Perfect Scores"	Interquartile Range
13,870	0.8%	2.7%	0.0%	0.0%	0.0%	1.0%	2.2%	1.0%	13,870

Analysis date: 7/3/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting of this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (.\quarter_4_5\db178_request\db178_request_017.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 long-stay residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examined the percentage of nursing homes that can report each measure (referred to as the QM reportability). ***Table 4.9-4*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 88.4 percent of facilities had sufficient sample size to meet minimum requirements for public reporting of QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) using MDS 3.0. This is comparable to the MDS 2.0 measure, which 88.5 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. Therefore, the reportability of short-stay QMs may be substantially improved after including the discharge assessment, but the impact on the long-stay QM should be less marked.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay), we examined changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.9-4***. After excluding discharge

assessments from the target assessment sample, reportability was completely unchanged for QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) (88.4 percent with discharge assessments and without discharge assessments).

Table 4.9-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0683 Percent of Long-stay Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator	1,203,726 (100%) [a]	1,202,329 (99.9%) [a]
Number (%) of resident episodes in the numerator	1,125,833 (93.5%) [a]	1,124,391 (93.4%) [a]
Number (%) of facilities able to report this QM	13,870 (88.4%) [b]	13,862 (88.4%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay sample (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\ )

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.9-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. For each pair of quarters, rankings were extraordinarily stable: at least four-fifths of facilities stayed in approximately the same rank, and very few (no more than 2.6 percent) facilities saw their rank change by more than three deciles.

Table 4.9-5
Distribution of change in facility rank from one quarter to the next, QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	11,866 (87.2%)	11,725 (85.4%)	11,043 (80.4%)
Between 1 and 2 deciles	1,236 (9.1%)	1,289 (9.4%)	1,662 (12.1%)
Between 2 and 3 deciles	393 (2.9%)	527 (3.8%)	674 (4.9%)
More than 3 deciles	117 (0.9%)	183 (1.3%)	353 (2.6%)
Total	13,612	13,724	13,732

NOTES:

Total number of facilities in each column reflects all the facilities that could report the measure in both quarters.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 through Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.9-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 12.4 percent). Scores for this measure changed little from Quarter 3 of 2011 to Quarter 4. Overall, the average change in scores for this measure was an increase of 0.1 percentage point. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (88.6 percent saw changes of less than one standard deviation), whereas just over 2 percent saw changes of more than three standard deviations.

Table 4.9-6
Changes in facility scores from one quarter to the next, QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
13,732	0.1%	6.0%	0.9%	1.0%	4.0%	61.2%	27.4%	3.2%	1.1%	1.2%

NOTES:

Number of facilities reporting (13,732) reflects facilities that reported this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures, should be correlated). Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures where applicable. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) and the facility's percentile rank on QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay), on QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay), and on QM #0682 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Short Stay) for Quarter 4, 2011(as these measures are all concerned with the vaccination process). We found that the correlation between the nursing home's percentile rank on QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) and percentile rank on QM #0680 Percent of Residents Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) was moderate (correlation coefficient = .560) and statistically significant ($p < 0.001$).[25] The correlation between rank on QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) and QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) was moderate (r =.392) and statistically significant ($p < 0.001$), and the correlation between rank on QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) and QM #0681 Percent of Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay) was high ($r = 0.690$) and statistically significant ($p < 0.001$).

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether State characteristics might be a source of facility score variation for QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) and thus a potential threat to the measure's validity, we conducted a one-way analysis of variance (ANOVA). We examined the interquartile difference between the mean State-level scores across States. The proportion of variance in this measure explained by the State in which facilities are located is 4.5 percent and

[25] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

significant [$F(51, 13{,}818) = 12.7$, $p < .001$].[26] The difference between the mean State-level scores for States at the 25th percentile and the 75th percentile is 3.6 percent. Thus, although the majority of the variance in QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) scores is due to factors other than geography, a small and significant proportion of the variance is explained by the respective States in which nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. There are no exclusion criteria for this measure, and values of [0] are imputed for missing data items. Thus, all 1,203,726 long-stay residents were included in the construction of this quality measure.

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2) the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.9-7.*** For this measure, virtually no facilities are affected by excluding discharge assessments from the calculation.

Table 4.9-7
Distribution of facility-level score change after excluding discharge assessment for QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
13,862	0.0%	0.9%	0.0%	0.0%	0.0%	0.0%	0.0%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request_v1\db135_request_v1.log)

Seasonal Variation

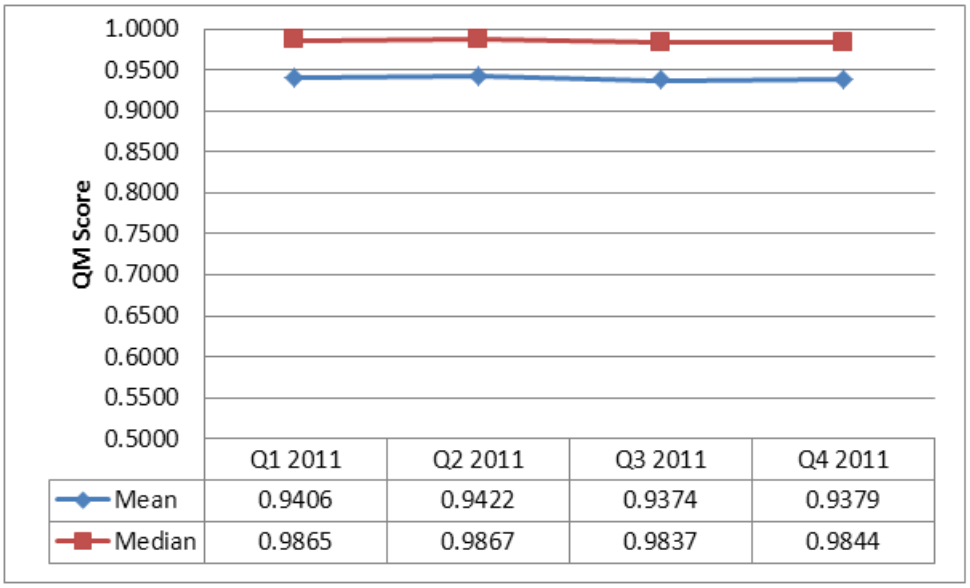
Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address this interest in

[26] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

whether seasonal variation might play a role in the score for QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay), we examined the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.9-1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, was 94.1 percent, 94.2 percent, 93.7 percent, and 93.8 percent. The quarterly national median was 98.7 percent, 98.7 percent, 98.4 percent, and 98.4 percent. Both mean and median remained quite high across all quarters, fluctuating up and down very little. With only 12 months of data available, it is too early to ascertain if the changes in national QM score over time are consistent with seasonal variation, this will need to be confirmed when multiple years of data become available.

Figure 4.9-1
Seasonal (quarterly) variation in QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay)



SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 - Quarter 4, 2011 \qm_quarter_1_2\complete\nh_017_10.log f \qm_quarter_2_3\ complete\nh_017_10.log \qm_quarter_3_4\ complete\nh_017_10.log \qm_quarter_4_5\ complete\nh_017_10.log

4.9.5 Risk Adjustment

QM #0683 Percent of Residents Who Were Assessed and Appropriately Given the Pneumococcal Vaccination (Long-Stay) was endorsed by NQF without denominator exclusion and model-based risk adjustment.

4.10 QM# 0684 Percent of Residents With a Urinary Tract Infection (Long Stay)

4.10.1 Summary of Findings

QM# 0684 Percent of Residents With a Urinary Tract Infection (Long Stay) reports the percentage of long-stay residents whose target assessment indicating urinary tract infection (UTI) within the last 30 days. This subsection presents descriptive analyses and test results for this QM. We found that 98.1 percent of long-stay resident episodes meet the denominator inclusion criteria and thus are included in the denominator. Based on these resident episodes, 87.7 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 7.7 percent. The QM score varies across facility with a standard deviation of 5.7 percent and interquartile range of 7.3 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

As reliability checks of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. About half of facilities remained within the same decile from quarter to quarter. About 90 percent of facilities are with rank changes within three deciles. These findings indicate that this QM has good reliability.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures because they reflect similar care processes. We selected QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) and QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the three QMs. Second, we examined the frequency and distribution of missing data. Missing rate is very low for the items used to construct this QM and therefore should not pose a threat to validity of the QM. Lastly, we analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 3.4 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

4.10.2 Background and Introduction to QM

This QM measures the percentage of long-stay residents who have a target MDS assessment that indicates a UTI within the last 30 days. In order to address seasonal variation, the proposed measure uses a 6-month average for the facility. Long-stay nursing facility residents are those whose cumulative stay in the facility is over 100 days. Target assessments may be discharge assessments; PPS 14-, 30-, 60-, or 90-day assessments; or quarterly, annual, significant change, or significant correction OBRA assessments. Admission (A0310A = 1), 5-day PPS and readmission/return assessments (A0310B = 1, 6) are excluded from the sample because UTI s identified on these assessments may have been acquired in other settings than the nursing facility.

A. Background for this QM

Nursing facility residents often develop infections, (Finnegan, Austin, & Cape, 1985; Jackson, Fierer, Barrett-Conner, & et al., 1992; Magaziner, Tenney, Deforge, & et al., 1991; Nicolle, McIntyre, Zacharies, & et al., 1984; Strausbaugh & Joseph, 2000) and among these, UTIs are the most common (Katz, Beam, Brand, & et al., 1990; Lee, Thrupp, Friis, & et al., 1992; Zimmer, Bentley, Valenti, & et al., 1986). A multistate study found an incidence rate of 45.3 persons hospitalized for UTI per 1,000 person-years of observation (Rogers, Fries, et al., 2008; Rogers, Mody, et al., 2008). Some residents who develop UTIs develop blood infections, and 10 percent of these patients die within a week (Saint, Kauman, Robers, & et al., 2006). Symptoms of UTIs include fever, painful or difficult urination, increased frequency and urgency of urination, blood in the urine, low abdominal or flank pain or tenderness, and deterioration in mental status (such as increased confusion). Using MDS 2.0 data for July–September 2010, the national prevalence of UTIs in nursing facilities was 9.7 percent, with a range from a low average of 5.7 percent in Hawaii to a high average of 13.2 percent in West Virginia (Centers for Medicare & Medicaid Services, 2009, 2010c). The UTI quality measure is the only measure in the current measure set that addresses infections, making it an important indicator of how facilities prevent and manage infections.

In a clinical review of the nursing home QMs using the MDS 2.0, a Technical Expert Panel (TEP) concluded that the urinary tract infection quality measure is a "valuable source of information for nursing homes" (Brega, Goodrich, Nuccio, et al., 2008). The measure prompts facilities to examine their approach to perineal care and their general infection rate. These infections have the potential for significant morbidity and mortality (Nicolle, 1993; Nicolle & SHEA Long-Term-Care Committee, 2001). Infections increase the use of medical care and costs. Because many UTIs are related to catheters, this quality measure provides an additional incentive for the facility to monitor its catheter use (Gould et al., 2009).

Some UTIs can be prevented by keeping the periurethral area clean, emptying the bladder regularly, drinking enough fluids, and practicing good hygiene (J. Gammack, 2003). Finding the cause and getting early treatment of a urinary tract infection can prevent the infection from spreading and becoming more serious or causing complications, such as delirium. Urinary incontinence, catheter use and UTIs are interrelated. Indwelling urinary catheterization can frequently cause UTIs in the elderly. Catheters are commonly used to address incontinence. When not properly maintained and monitored, indwelling catheters can cause chronic pain or infection leading to a greater functional decline and decreased quality of life for the resident (J. Gammack, 2003). At least 40 percent of all infections seen in nursing homes are in the urinary tract system; of those infections, 80 percent are due to urinary tract catheterization and instrumentation (Newman, Fader, & Bliss, 2004). Because of the interrelationship between UTI and catheter use, and the NQF has recommended that the long-stay UTI QM and long-stay catheter QM be presented as paired measures.

Although research suggests racial disparities in quality of care in nursing homes between African Americans and whites (Grabowski, 2004; Howard et al., 2002; Miller, Papandonatos, Fennell, & Mor, 2006; Mor et al., 2004; Smith, Feng, Fennell, Zinn, & Mor, 2007) and between Hispanics and whites (Fennell, Feng, Clark, & Mor, 2010), past research has not specifically

addressed racial disparities in UTIs. Furthermore, no past research has examined other types of disparities (e.g., ethnicity, rural/urban, or income) specifically for this measure.

B. Summary of Differences Between MDS 2.0 to MDS 3.0 QM definitions

The underlying data items used to calculate the MDS 3.0 measure remain unchanged from MDS 2.0 to the MDS 3.0 assessment. However, the proposed measure modifies the MDS 2.0 quality measure by using a 6-month average for the facility rather than data from just one quarter to address concerns about fluctuation in QM scores attributable to seasonal variation in resident risk for UTI. Unpublished data analysis of MDS 2.0 by Dr. Vincent Mor at Brown University found seasonal variation in the rate of urinary tract infections in nursing homes.

Additionally, the MDS 3.0 includes more detailed instructions for identifying UTIs. Instructions were designed with input from infectious disease experts from CDC and included more explicit instructions for clinicians to determine whether the diagnosis was active (Saliba & Buchanan, 2008). There were some concerns regarding the validity and reliability of the MDS 2.0 measure related to a substantial number of false positives. To address this issue, the MDS 3.0 manual provides more detailed instructions for identifying urinary tract infections, in particular, all four criteria listed in the RAI Manual, Chapter 3, Section I, Coding Tips, must be met for the resident to be considered to have a urinary tract infection (Centers for Medicare & Medicaid Services, 2010b). The criteria are as follows:

1. a diagnosis of UTI in the last 30 days by a physician, nurse practitioner, physician assistant, or clinical nurse specialist or other authorized licensed staff as permitted by state law;
2. sign or symptom attributed to UTI, which may include fever, urinary symptoms (e.g., peri-urethral site burning sensation, frequent urination of small amounts), pain or tenderness in flank, confusion or change in mental status, change in character of urine (e.g., pyuria);
3. "significant laboratory findings" (The attending physician should determine the level of significant laboratory findings and whether or not a culture should be obtained); and
4. current medication or treatment for a UTI in the last 30 days.

Initial reliability testing of the MDS 3.0 items resulted in substantial kappas of 0.71 for the UTI item (I2300), however these kappas were lower than many of the other items evaluated by interrater reliability testing (Saliba & Buchanan, 2008). Not surprisingly facility level scores using MDS 2.0 and MDS 3.0 QM definitions were highly correlated (0.80) (Saliba & Buchanan, 2008).

In summary, the items used for the MDS 3.0 long-stay UTI QM and the calculation of the numerator and denominator are the same as for the MDS 2.0; however measure improvements included more detailed instructions to address false positive rates; measure specification based on a 6-month average to address seasonal variation; and a change in the QM sample definition from chronic care where specification was based on just a full or quarterly MDS in the target quarter. For the MDS 3.0, the analogous sample long-stay residents are defined as residents with more than 100 cumulative days in facility and the QM can be based on any PPS or OBRA

assessment except for an unscheduled assessment used for PPS (A0310B = 7), with the exception of a few QMs, which may exclude certain assessments, like the exclusion of admission, 5-day PPS and readmission/return assessments for the UTI QM.

C. Summary of Analyses

The following sections summarize the basic analyses performed on this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- Number of assessments that are included and excluded from the numerator of the QM based on sample restrictions
- Whole sample prevalence of items that are related to the QM
- Findings regarding variability, reportability, reliability, and validity

4.10.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, a total of 1,180,633 episodes (98.1 percent) were included in the denominator of QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay), and 23,093 resident episodes (1.9 percent) of the total number of long-stay resident episodes) were excluded from the denominator. ***Table 4.10-1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criteria. Note that one resident episode can meet multiple exclusion criteria. The number and percent of resident episodes that met each exclusion criterion are as following:

- A total of 701 (0.0 percent) assessments met exclusion criterion 1: missing data on the urinary tract infection item; and
- 22,402 (1.9 percent) assessments met exclusion criterion 2: a target assessment that was an admission assessment or a PPS 5-day or readmission/return assessment.

Table 4.10-1
Target assessments used in the calculation of QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	1,180,633	98.1%
Excluded—Exclusion 1	701	0.0%
Excluded—Exclusion 2	22,402	1.9%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Missing Data
Exclusion 2 = Wrong target assessment (admission, PPS 5-day or readmission/return assessment)

[a] Column values may not add up to total since a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

The numerator for the UTI QM is based on a single item, which asks whether there was an active UTI diagnosis in the last 30 days prior to the assessment. ***Table 4.10-2*** shows the distribution of responses on this item for the total set of long-stay target assessments including those excluded from the QM sample (1) for incomplete data and (2) because the target assessment was an admission, PPS 5-day, or readmission/return assessment. Overall, 7.2 percent of the more than 1 million responses triggered the event, indicating a UTI was present. The proportion of UTIs excluded from the QM sample because the target assessment was the wrong type is quite small (22,392, or 1.9 percent); exclusions due to missing responses were even less frequent (701, or 0.0 percent).

Table 4.10-2
Frequency and percentage responses on items used to define the numerator for QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay)

Response	Included *n*	Included %	Excluded *n*	Excluded %	Total
No response	0	0.0%	701	0.0%	701
No UTI	1,092,924	90.8%	16,797	1.4%	1,109,721
UTI	87,709	7.2%	5,595	0.5%	93,304
Total	1,180,633	98.1%	23,093	1.9%	1,203,726

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_018_10.log)

4.10.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles) or shown by the percentage of nursing homes with "perfect scores" (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes), which may indicate a "ceiling effect." ***Table 4.10-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) is 7.7 percent with a standard deviation of 5.7 percent. This QM has an interquartile range of 7.3 percent. About 7.2 percent of facilities have scores of 0 percent ("perfect" scores).

Table 4.10-3
QM score distribution for QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay)

n	Mean score	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
13,773	7.7%	5.7%	1.3%	3.5%	6.7%	10.8%	15.2%	7.2%	7.3%

Analysis date: 5/29/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_018_10.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examine the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.10-4*** shows the results of this analysis using of MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 87.7 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) using MDS 3.0. This is a slight improvement over the MDS 2.0 measure, which 85.6 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, slightly more than 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and that 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may improve the reportability of long-stay QMs to some extent. Therefore, the reportability of short-stay QMs may be substantially improved after including the discharge assessment, but the impact on the long-stay QM should be less marked.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay), we report changes in the number of facilities able to report this QM when

discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.10-4.*** After excluding discharge assessments from the target assessment sample, reportability is almost completely unchanged for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) (87.73 percent with discharge assessments and 87.71 percent without discharge assessments).

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.10-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. For each pair of quarters, facility rankings were relatively variable. Just over half of the facilities reporting on this measure maintained their rank from each quarter to the next, and regularly changed more than 2 deciles (more than 20 percent of facilities in each quarter-to-quarter transition). However, these ranking shifts, combined with the analysis of facility scores, described subsequently, indicate that small shifts in QM scores for this measure can lead to significant shifts in relative rank.

Table 4.10-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay)

Key Value	Discharge Assessment Included	Discharge Assessment Excluded
Number (Percent) of resident episodes in the denominator	1,180,633 (98.1%) [a]	1,174,658 (97.6%) [a]
Number (Percent) of resident episodes in the numerator	87,709 (7.3%) [a]	86,832 (7.2%) [a]
Number (Percent) of facilities able to report this QM	13,760 (87.7%) [b]	13,758 (87.7%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay population (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one "short-stay" or long-stay residents (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Table 4.10-5
Distribution of change in facility rank from one quarter to the next for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	7,406 (54.9%)	7,527 (55.3%)	7,541 (55.4%)
Between 1 and 2 deciles	2,696 (20.0%)	2,774 (20.4%)	2,671 (19.6%)
Between 2 and 3 deciles	1,602 (11.9%)	1,562 (11.5%)	1,610 (11.8%)
More than 3 deciles	1,788 (13.3%)	1,741 (12.8%)	1,798 (13.2%)
Total	13,492	13,604	13,620

NOTES:

Total *n*'s reflect facilities that meet minimum requirements for public reporting this QM in both Quarters.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 through Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.10-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 5.7 percent). Overall, the average change in scores for this measure was a decline of 0.1 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (73.7 percent saw changes of less than one standard deviation), while approximately 1 percent saw changes of more than three standard deviations.

Correlations. A common strategy used to evaluate convergent validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes. Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures where applicable. Among nursing homes that could report both related measures, we calculated the correlations between the facility's percentile rank on QM# 0684 Percent of Residents With a Urinary Tract Infection (Long Stay) and the facility's percentile rank on QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) and on QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) in Quarter 4, 2011 (because the three all measure processes and/or

Table 4.10-6
Change in facility score from one quarter to the next for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
13,620	-0.1%	4.5%	0.5%	2.4%	10.3%	33.8%	39.9%	10.2%	2.4%	0.5%

NOTES:

Number of facilities reporting (13,620) reflects facilities that meet minimum requirements for public reporting this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

outcomes related to the health of the urinary tract). We found that the correlations between the nursing home's percentile rank on QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) and percentile rank on QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) were both small ($r = 0.092$ and $r = 0.117$, respectively) and statistically significant ($p < 0.001$ in both cases).[27] The statistically significant correlations between these measures to some extent support reporting QM# 0684 Percent of Residents With a Urinary Tract Infection (Long Stay) as a paired measure.

Validity

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as state-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether state characteristics might be a source of facility score variation for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) and thus a potential threat to the measure's validity we conducted a one-way analysis of variance (ANOVA). We also did a simple examination of the interquartile difference between the mean state-level scores across states. The proportion of variance in this measure explained by the state in which facilities are located is 3.4 percent and significant [$F(51, 13{,}721) = 10.8$, $p < .001$].[28] The difference between the mean state-level scores for states at the 25th percentile and the 75th percentile is 1.7 percentage points. Thus, while the majority of the variance in QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) scores is due to factors other than geography, there is a small and significant proportion of the variance that is explained by the respective states in which nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns indicate that certain types of residents tend to have assessments with missing data in ways that impact the calculation of a quality measure, then that measure may be not be capturing processes and outcomes for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

Facility-level analysis of MDS 3.0 data from the fourth quarter of 2011 indicates that missing data on items used to construct this measure were not prevalent and thus unlikely to threaten the validity of this measure. After excluding those residents with admission, or PPS 5-day or readmission/return assessments (22,402 residents, or 1.9 percent of the long-stay population), fewer than one-tenth of 1 percent of residents (701) were excluded due to missing data on the item used to indicate UTI (I2300).

---

[27] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

[28] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

Table 4.10-7 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. Missing data on items used to construct this quality measure are exceedingly rare: at least 90 percent of facilities have no missing data, and the mean rate is just 0.1 percent.

We also examined the relationship between missing data and QM scores. ***Table 4.10-8*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure; they are essentially identical and near-zero across quartiles. Further, there is no correlation between missing data and QM scores ($r = 0.01$, *n.s.*).

Table 4.10-7
Distribution of facility-level missing rate for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay)

n	Mean	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,399	0.1%	1.4%	0%	0%	0%	0%	0%

NOTES:

n = number of facilities who have data for the numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request.log)

Table 4.10-8
Distribution of facility-level missing rate for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) by quartile of QM score

Quartile of QM Score	*n*	Mean	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,833	0.1%	1.6%	0%	0%	0%	0%	0%
26%–50%	3,848	0.0%	0.4%	0%	0%	0%	0%	0%
51%–75%	3,932	0.1%	1.2%	0%	0%	0%	0%	0%
76%–100%	3,719	0.1%	1.0%	0%	0%	0%	0%	0%

NOTES:

n = number of facilities who have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in one, the overall impact of including the discharge assessment on QM rates and two,

the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the discharge assessment on QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.10-9.*** The mean nursing home level QM score and distributions are essentially unaltered (the mean change is less than one-tenth of 1 percent) after excluding discharge assessments from the set of eligible target assessments compared to including the discharge assessment. Thus, the use of the discharge assessment in the QM score calculations does not meaningfully alter the distribution of scores among facilities.

Table 4.10-9
Distribution of facility-level score change after excluding discharge assessment for QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay)

n	Mean score change	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
13,758	0.0%	1.1%	-1.3%	0.0%	0.0%	0.1%	1.2%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments

Analysis date: 5/29/2012

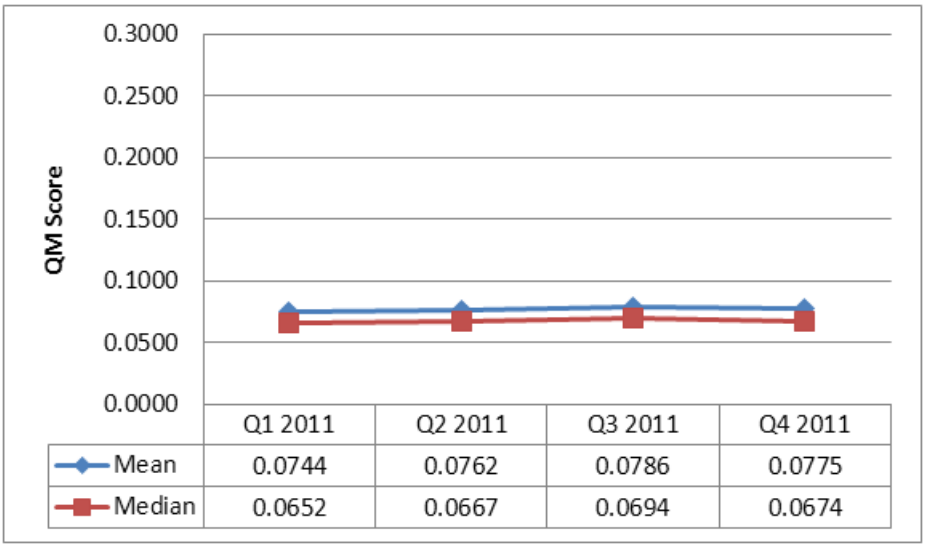
SOURCE: RTI analysis of MDS 3.0 data for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this potentially indicates a threat to measure validity as it may be influenced by factors outside of the nursing home's control. To address whether seasonal variation might play a role in the score for QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay), we examined the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.10-1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, was 7.4 percent, 7.6 percent, 7.9 percent, and 7.8 percent. The quarterly national median was 6.5 percent, 6.7 percent, 6.9 percent, and 6.7 percent. Both mean and median increased slightly from Quarter 1 2011 to Quarter 3 2011, and decreased slightly from Quarter 3, 2011, to Quarter 4, 2011. With only 12 months of data available, it is too early to ascertain if the changes in national QM score over time are consistent with seasonal variation, this will need to be confirmed when multiple years of data become available.

Figure 4.10-1
Seasonal (quarterly) variation in QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay)



SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 - Quarter 4, 2011 \qm_quarter_1_2\complete\nh_018_10.log f \qm_quarter_2_3\ complete\nh_018_10.log \qm_quarter_3_4\ complete\nh_018_10.log \qm_quarter_4_5\ complete\nh_018_10.log

4.10.5 Risk Adjustment

QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) was endorsed by NQF without denominator exclusion and model-based risk adjustment. During the development of the MDS 3.0 measure no major conditions were identified that were appropriate for risk adjustment and clearly associated with UTI (NQF, 2011).

4.11 QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)

4.11.1 Summary of Findings

QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) reports the percentage of low-risk long-stay residents whose target assessment indicating frequently or always incontinence of the bladder. This QM is risk adjusted, applying denominator exclusions. Residents at high-risk for incontinence are excluded from the denominator of this quality measure. High risk is defined as any of the following conditions: severe cognitive impairment; totally dependent in bed mobility, transfer, or locomotion.

This subsection presents descriptive analyses and test results for this QM. 39.7 percent of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 42.7 percent of the facilities are able to report this QM,(i.e., facilities with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 42.8 percent. The QM score varies across facility with a standard deviation of 17.3 percent and interquartile range of 24.7 percent,

indicating acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

To measure, we assessed changes in QM score and in rank based on the QM score for by facility by quarter. The majority of facility score changes, (either improvements or declines) were within one standard deviation. A very low proportion of changes being greater than three standard deviations. A similar pattern was found for facility rank changes. Most facilities (about 90 percent) illustrate rank changes within three deciles.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures since the measures reflect similar care processes. We selected QM# 0684 Percent of Residents With a Urinary Tract Infection (Long Stay) and QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the three QMs. Second, we examined the frequency and distribution of missing data. On average, 1 percent of resident episodes cannot be used for calculating this QM due to missing data. Missing data, therefore, should not pose a threat to validity of the QM. Last but not the least, we also analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 16.2 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

4.11.2 Background and Introduction to QM

This measure reports the percent of long-stay residents who are frequently or almost always bladder or bowel incontinent as indicated on the target MDS assessment during the selected quarter (3-month period). The proposed measure is stratified into high- and low-risk groups; only the percentage of the low-risk group (e.g., residents whose mobility and cognition are not impaired) is calculated and included as a publicly reported quality measure.

A. Background for this QM

At least 17 million Americans have urinary incontinence (UI). It is the second leading cause of institutionalization of the elderly, and occurs in over 50 percent of nursing home residents, with prevalence ranging from 43 percent to 77 percent (median 58 percent) (Lekan-Rutledge & Colling, 2003; Offermans, Du Moulin, Hamers, Dassen, & Halfens, 2009). It is important to treat as its prevention may reduce the likelihood of infections, pressure ulcers, and other health complications from poor health hygiene. Using MDS 2.0 data, prevalence of positive response to the item Occasional or Frequent Bladder Incontinence Without a Toileting Plan in nursing homes in the third quarter of 2010 was estimated to be between 23.6 percent in Wisconsin and 84.2 percent in Colorado (Centers for Medicare & Medicaid Services, 2006). For the same quarter, the results of the QM Low-Risk Residents Who Lose Control of Their Bowels or Bladder (based on MDS 2.0 data) ranged from 38.2 percent to 69.0 percent, with a national average of 51.0 percent (Centers for Medicare & Medicaid Services).

The impact of incontinence profoundly affects nursing home residents as well as staff. Incontinence can cause feelings of shame and embarrassment for the resident and increases the burden of care for caregivers. General health and quality of life factors, such as emotional well-

being and social functioning, are also affected by incontinence. For both cognitively intact and cognitively impaired residents, bladder incontinence significantly predicted more dependence in total activities of daily living (ADLs), toileting, and personal hygiene (Wang, Chang, Eberly, Virnig, & Kane, 2010). Nursing home staff may view incontinence care as both difficult and burdensome. As a result, it is frequently managed inappropriately (Ouslander & Schnelle, 1995).

Loss of bowel and bladder control can be caused by:

- physical problems (e.g., constipation, muscle weakness, or a bladder infection);
- location problems (e.g., the bathroom is too far away);
- reaction to medication;
- limited ability to walk or move around;
- diet and fluid intake;
- toilet routine (e.g., timing trips to the bathroom);
- whether someone can provide assistance when needed; and
- certain medical conditions (e.g., residents with diabetes, dementia, spinal cord injury, or neurological disease are at a higher risk of losing bowel and bladder control) (Landefeld et al., 2008; MedQIC, 2002; Quality Measures Management Information System (QMIS).

Incontinence, particularly reversible conditions of incontinence, is treatable in many cases, and incontinence programs do make a difference. Nursing facility residents who are incontinent of urine should have a targeted physical examination, including a urinalysis and a determination of postvoid residual urine volume done by catheterization or ultrasonography (Ouslander & Schnelle, 1995). Scheduled toileting and bladder programs can be successfully implemented among nursing home residents. The key to the success of these programs is to appropriately identify residents who should be targeted for each specific program (Newman & Palmer, 1999). As with urinary incontinence, fecal incontinence may also be caused by potentially reversible conditions. After such conditions are excluded, fecal incontinence can generally be managed effectively by avoiding fecal impaction and by using a systematic bowel-training protocol (Ouslander & Schnelle, 1995).

A randomized controlled trial demonstrated that a multicomponent intervention showed significant improvement over the control group in urinary incontinence, number of bowel movements, and percentage of bowel movements in the toilet (Schnelle et al., 2010). The multicomponent approach addressed several of the factors thought to lead to urinary and fecal incontinence, such as inadequate food and fluid intake, low levels of physical activity, undetected issues with anorectal sphincter control, and cognitive barriers. The study found that residents with higher cognitive impairment responded better to the intervention than did residents with lower levels of cognitive impairment (Schnelle et al., 2010). Determining the cause of and initiating treatment for problems with bowel and bladder control are important for many reasons. Physically, managing bowel and bladder control can help prevent infections, pressure ulcers, and other complications from poor health hygiene. Treatment can promote the well-being of the resident by restoring dignity and social interaction.

Severe cognitive impairment is a significant risk factor for incontinence, and renders residents unable to participate in many interventions to reduce incontinence (Leung & Schnelle, 2008). Therefore, residents with severe cognitive impairment (as determined by the Brief Interview for Mental Status, or BIMS) are excluded from the measure. BIMS was tested and showed a sensitivity of 69.7 percent to 94.4 percent, and a specificity of 85.6 percent to 78.6 percent for severe cognitive impairment, indicating that the majority of patients with cognitive impairments are being correctly identified (Chodosh et al., 2008). The BIMS showed high reliability as well, with facility nurses and research assistants obtaining the identical mean scores (Chodosh et al., 2008).

B. Summary of Differences between MDS 2.0 and MDS 3.0 QM Definitions

In the MDS 3.0, 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 are incontinent of bowel (H0300 = 2 or 3) or bladder (H0400 = 2 or 3); where H0300 = 2 = frequently incontinent (seven or more episodes of bowel incontinence, but at least one episode of continent voiding), H0300 = 3 = always incontinent (no episodes of continent voiding), H0400 = 2 = frequently incontinent (two or more episodes of bowel incontinence, but at least one continent bowel movement), and H0400 = 3 = Always incontinent (no episodes of continent bowel movements).

The MDS 2.0 and 3.0 have different sets of response options and look-back windows for the items. The response options on MDS 2.0 items are (for urinary and bowel): continent, usually continent, occasionally incontinent, frequently incontinent, and incontinent in the last 14 days. The MDS 3.0 refers to episodes over the last 7 days, with response options of: always continent, occasionally incontinent, frequently incontinent, always incontinent or not rated due to use of catheter, and ostomy or lack of output or bowel movement for the entire 7 days. The MDS 3.0 eliminates the "usually continent" frequency rating to simplify response categories.

Level of cognitive function measures, which are also used to identify high risk of incontinence incorporated into the exclusion criteria, differ between the 2.0 and 3.0. The items in the 2.0 are all based on staff assessment. The items in the MDS 3.0 include a direct assessment of cognitive function (the BIMS) and staff assessment. The number of residents identified as severely cognitively impaired (and hence to be excluded) is much higher using the BIMS than using the MDS 2.0 measure and the logic methodology to identify cognitive impairment is more complex.

For both the 2.0 and the 3.0 version of the measure, the denominator includes all residents with target assessment, except those with exclusions. Both the MDS 2.0 and MDS 3.0 versions of the measure exclude residents from the denominator who are who are new admissions or readmissions; who are missing data for the incontinence item; or who are comatose, have an indwelling catheter, or have an ostomy. They also exclude residents who are high risk, defined as those who have severe cognitive impairment, total dependence in bed mobility, transfer self-performance, and locomotion on unit self-performance. They also exclude residents who are not already in the numerator and are missing data for any high-risk items.

Note also that the overall sample definitions also changed from MDS 2.0 to MDS 3.0. For the MDS 2.0, residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments. For the MDS 3.0, the analogous sample is for long-stay residents, defined as residents with more than 100 cumulative days in facility, and assessments may be discharge assessments; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

The MDS 2.0 version of this measure performed poorly on tests of validity. Previous research indicates this measure was not well correlated with any of the other nursing home QMs. In the MDS 3.0 version, the incontinence measure showed small but significant correlations with the other urinary tract-related measures: $r = .1341$ ($p < .001$) for QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay); $r = .0502$ ($p < .001$) for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) (Brega et al., 2007).

C. Summary of Analyses

The following sections summarize the basic analyses performed on this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- Number of assessments that are included and excluded from the numerator of the QM based on sample restrictions
- Whole sample prevalence of items that are related to the QM
- Findings regarding variability, reportability, reliability, and validity

4.11.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 477,004 (39.7 percent) assessments included in the denominator of QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay), and a total of 726,722 resident episodes (60.4 percent of the total number of long-stay resident episodes) were excluded from the denominator. ***Table 4-11.1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criteria. Note that one resident episode can meet multiple exclusion criteria. Overall, the number and percent of resident episodes that met each exclusion criterion are as follows:

- a total of 14,414 (1.2 percent) assessments met exclusion criterion 1: missing data on items used to construct the numerator;
- 22,402 (1.9 percent) met exclusion criterion 2: the assessment was an admission assessment or a PPS 5-day or readmission/return assessment;

- 679,279 (56.4 percent) assessments met exclusion criterion 3: presence of high-risk resident characteristics (severe cognitive impairment or total dependence in bed mobility, transfer, and/or locomotion);
- 4,554 (0.4 percent) assessments met exclusion criterion 4: comatose state;
- 69,365 (5.8 percent) of residents met exclusion criterion 5: presence of indwelling catheter or missing data for indwelling catheter status; and
- 22,009 (1.8 percent) met exclusion criterion 6: presence of indwelling catheter or missing data for ostomy status.

Table 4.11-1
Target assessments used in the calculation of QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	477,004	39.6%
Excluded—Exclusion 1	14,414	1.2%
Excluded—Exclusion 2	22,402	1.9%
Excluded—Exclusion 3	679,279	56.4%
Excluded—Exclusion 4	4,554	0.4%
Excluded—Exclusion 5	69,365	5.8%
Excluded—Exclusion 6	22,009	1.8%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Missing data
Exclusion 2 = Wrong target assessment type (admission assessment, PPS 5-day or readmission/return assessment)
Exclusion 3 = Residents with high-risk conditions
Exclusion 4 = Comatose
Exclusion 5 = Presence of indwelling catheter or missing data for indwelling catheter status
Exclusion 6 = Presence of ostomy or missing data for ostomy status

[a] Column values may not add up to total since a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident assessments in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of 2011 MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items – Distribution of Triggering Events

Table 4.11-2 describes the responses to H0300 (Bladder Incontinence) and H0400 (Bowel Incontinence), the MDS 3.0 items used to calculate this QM, for those individuals included in the calculation of this QM. Incontinence was common among the long-stay sample: 59.4 percent of target assessments indicated frequent or persistent bladder incontinence (conditions that trigger numerator inclusion: H0300 = [2,3]) and 50.5 percent of target assessments indicated frequent or persistent bowel incontinence.

Table 4.11-2
Responses to H0300 (Bowel Incontinence) and H0400 (Bladder Incontinence) on target assessments not otherwise excluded from analysis

Item	Entry	Frequency	Percentage
H0300	- = Missing	1,807	0.01%
H0300	0 = Always continent	259,647	21.6%
H0300	1 = Occasionally incontinent (< 7 episodes of incontinence)	161,637	13.4%
H0300	2 = Frequently incontinent (7 or more episodes of urinary incontinence, but at least one episode of continent voiding)	295,892	24.6%
H0300	3 = Always incontinent (no episodes of continent voiding)	417,965	34.3%
H0300	9 = Not rated, resident had a catheter, urinary ostomy, or no urine output for entire 7 days	66,778	5.5%
H0300	Total	1,203,726	—
H0400	- = Missing	1,730	0.0%
H0400	0 = Always continent	448,827	37.3%
H0400	1 = Occasionally incontinent (one episode of bowel incontinence)	127,685	10.6%
H0400	2 = Frequently incontinent (2 or more episodes of bowel incontinence, but at least one continent bowel movement)	186,388	15.5%
H0400	3 = Always incontinent (no episodes of continent bowel movements)	419,945	34.9%
H0400	9 = Not rated, resident had an ostomy or did not have a bowel movement for the entire 7 days	19,151	1.6%
H0400	Total	1,203,726	—

NOTE: H0300 = Urinary Continence; H0400 = Bowel Continence

Analysis date: 5/29/2012

SOURCE: RTI analysis of 2011 MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_019_10.log)

4.11.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with "perfect scores" (i.e., 0 percent triggering rates for

measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes), which may indicate a "ceiling effect." ***Table 4.11-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0685 is 42.8 percent with a standard deviation of 17.3 percent. This QM has an interquartile range of 24.7 percent. About 0.3 percent of facilities have scores of 0 percent ("perfect" scores).

Table 4.11-3
QM score distribution for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)

n	Mean score	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
6,689	42.8%	17.3%	20.3%	30.4%	42.6%	55.1%	65.2%	0.3%	24.7%

Analysis date: 5/29/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level.

SOURCE: RTI analysis of 2011 MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_019_10.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examine the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.11-4*** shows the results of this analysis using of MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 41.7 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) using MDS 3.0. This is a decline from the MDS 2.0 measure, which 72.4 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30- day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011 were discharge assessments, and that 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may improve the reportability of

long-stay QMs to some extent. Therefore, the reportability of short-stay QMs may be substantially improved after including the discharge assessment, but the impact on the long-stay QM should be less marked.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay), we report changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.11-4***. After excluding discharge assessments from the target assessment sample, reportability is largely unchanged for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) (41.7 percent with discharge assessments vs. 42.7 percent without discharge assessments).

Table 4.11-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (Percent) of resident episodes in the denominator	477,004 (39.6%) [a]	477,579 (39.7%) [a]
Number (Percent) of resident episodes in the numerator	197,852 (16.4%) [a]	197,443 (16.4%) [a]
Number (Percent) of facilities able to report this QM	6,545 (41.7%) [b]	6,700 (42.7%)

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay population (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one "short-stay" or long-stay residents (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of 2011 MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.11-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. Between Quarters 1 and 2 and again between Quarters 2 and 3, facility rank for this measure was exceptionally stable, with at least 76.6 percent of facilities maintaining the same rank within one decile, and 2 percent or fewer changing by more than three deciles. From Quarter 2 to Quarter 3, facility rankings were somewhat less stable: although the majority of facilities maintained the same ranking, 11.3 percent changed by more than 30 percentile points.

Table 4.11-5
Distribution of change in facility rank from one quarter to the next for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	4,119 (76.6%)	4,440 (78.1%)	7,852 (59.1%)
Between 1 and 2 deciles	866 (16.1%)	899 (15.8%)	2,439 (18.4%)
Between 2 and 3 deciles	287 (5.3%)	268 (4.7%)	1,503 (11.3%)
More than 3 deciles	106 (2.0%)	78 (1.4%)	1,495 (11.3%)
Total	5,378	5,685	13,289

NOTES:

Total *n*'s reflect facilities that meet minimum requirements for public reporting this QM in both Quarters.

Analysis date: 6/15/2012

SOURCE: RTI analysis of 2011 MDS 3.0 episode files for Quarter 1 through Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.11-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 17.3 percent). Overall, the average change in scores for this measure was an increase of 0.7 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (71.2 percent saw changes of less than one standard deviation), while fewer than 1 percent of facilities saw changes of more than three standard deviations.

Table 4.11-6
Change in facility score from one quarter to the next for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of Facilities That Declined by More than Three Standard Deviations:	Proportion of Facilities That Declined by between Two and Three Standard Deviations:	Proportion of Facilities That Declined by between One and Two Standard Deviations:	Proportion of Facilities That Declined by Less than One Standard Deviation	Proportion of Facilities That Improved by Less than One Standard Deviation	Proportion of Facilities That Improved by between One and Two Standard Deviations:	Proportion of Facilities That Improved by between Two and Three Standard Deviations:	Proportion of Facilities That Improved by More than Three Standard Deviations:
5,948	0.7%	7.2%	0.3%	2.0%	12.3%	36.2%	35.0%	11.1%	2.5%	0.5%

Analysis date:

SOURCE: RTI analysis of 2011 MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home QMs have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated). Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures, where applicable. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) and the facility's percentile ranks on QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay) and QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) (because these measures all concern processes and outcomes related to the health of the urinary tract) in Quarter 4, 2011. We found that the correlation between the nursing home's percentile rank on QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) and percentile rank on QM #0684 Percent of Residents with a Urinary Tract Infection (Long Stay) was small ($r = 0.092$) and statistically significant ($p < 0.001$). There was no significant correlation between percentile rank on QM #0685 on QM #0686 ($r = 0.017$, n.s.).[29]

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as state-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether state characteristics might be a source of facility score variation for QM #0685 and thus a potential threat to the measure's validity we conducted ANOVA. We also did a simple examination of the interquartile difference between the mean state-level scores across states. The proportion of variance in this measure explained by the state in which facilities are located is 16.2 percent and significant [$F(50, 6{,}638) = 25.6$ $p < .001$].[30] The difference between the mean state-level scores for states at the 25th percentile and the 75th percentile is 7.6 percentage points. Thus, while the majority of the variance in QM #0685 scores is due to factors other than geography, there is a relatively large (for the measures in this analysis) and significant proportion of the variance that is explained by the respective states in which nursing facilities are located. This feature of the measure warrants future monitoring.

---

[29] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

[30] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns indicate that similar types of residents and/or assessments have missing data it is likely that the measure may be not be capturing processes and outcomes for the intended population, thus inflating or suppressing facility level QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

In Quarter 4 of 2011, 726,722 residents (60.4 percent of the long-stay sample) were excluded from the construction of this measure. After excluding residents with admission assessments or PPS 5-day or readmission/return assessments (these were 22,402 residents: 1.9 percent of the long-stay population), the majority of residents (56.4 percent) were excluded because they did not meet the low-risk criteria for this measure (that is, they had cognitive impairment and/or major dependency on activities of daily living). Of the remaining residents, 4,554 (0.4 percent of the population) were excluded because they were comatose, 69,365 (5.8 percent) because they had an indwelling catheter, 22,009 (1.8 percent) because they had an ostomy, and 14,414 (1.2 percent) because of missing data on items used to construct this measure.

Table 4.11-7 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. Missing data are rare on the items used to construct this measure, with a facility-level mean missing rate of 1.2 percent. It is worthwhile to note that the small amount of missing data are concentrated in a few facilities: at the 90th percentile, the facility-level missing rate is 4.2 percent.

Table 4.11-7
Distribution of facility-level missing rate for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)

n	Mean	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,399	1.2%	3.2%	0%	0%	0%	1.2%	4.2%

NOTES:

n = number of facilities who have data for the numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request.log)

We also examined the possible relationship between missing data and QM scores. ***Table 4.11-8*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure. At each quartile, missing-data rates average slightly more than 1 percent, and the distributions are essentially identical. there is no evidence of statistically significant correlation between missing data and QM scores ($r = 0.005$, *n.s.*).

Table 4.11-8
Distribution of facility-level missing rate for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) by quartile of QM score

Quartile of QM score	*n*	Mean	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,829	1.3%	3.1%	0%	0%	0%	1.0%	4.3%
26%–50%	3,781	1.2%	2.7%	0%	0%	0%	1.3%	4.1%
51%–75%	3,809	1.2%	2.8%	0%	0%	0%	1.2%	4.2%
76%–100%	3,799	1.3%	2.9%	0%	0%	0%	1.4%	4.2%

NOTES:

n = number of facilities who have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in one, the overall impact of including the discharge assessment on QM rates and two, the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the discharge assessment on QM #0674 by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4-11.9.*** The mean nursing home–level QM score and distributions change slightly before and after excluding discharge assessments from the set of eligible target assessments. The average facility score decreased by 0.1 percent (with a median change of 0 percent) when discharge assessments are excluded. These findings demonstrate that including the discharge assessment in the QM score calculations does not meaningfully alter the distribution of scores among facilities.

Table 4.11-9
Distribution of facility-level score change after excluding discharge assessment for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)

n	Mean score change	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
6,541	-0.1%	1.8%	-2.3%	-1.0%	0.0%	0.7%	2.0%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Seasonal Variation

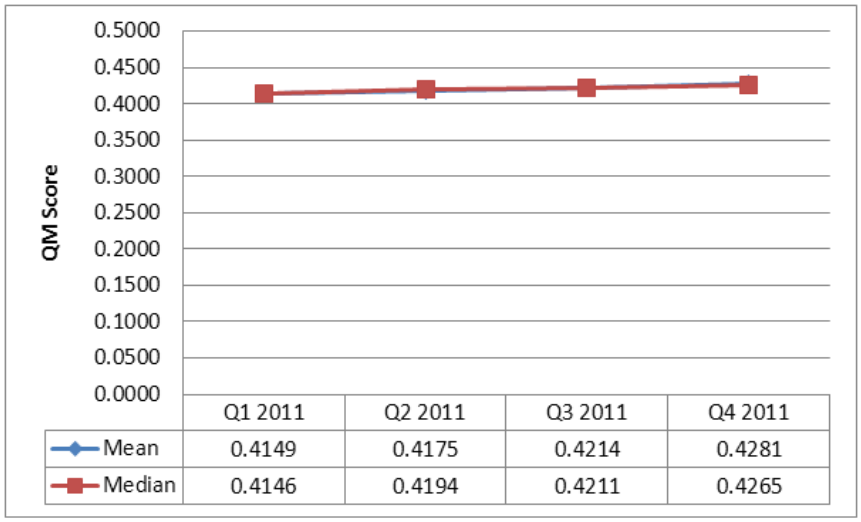
Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, the measure's validity is suspect and likely influenced by factors outside of the nursing home's control. To address potential seasonal variation on the score for QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay), we examine the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4-11.1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, was 41.5 percent, 41.8 percent, 42.1 percent, and 42.6 percent. The quarterly national median was 41.5 percent, 41.9 percent, 42.1 percent, and 42.5 percent. Both mean and median increased slowly over all four quarters. With only 12 months of data available, it is too early to ascertain if the changes in national QM score over time are consistent with seasonal variation, this will need to be confirmed when multiple years of data become available.

4.12.5 Risk Adjustment

QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay) is risk adjusted by applying denominator exclusions. Residents at high risk for incontinence are excluded from the denominator of this quality measure. High risk is defined as any of the following conditions: severe cognitive impairment; totally dependent in bed mobility, transfer, or locomotion. As presented previously in ***Table 4.11-1***, slightly more than half of the long-stay residents are excluded due to high risk. The denominator exclusions strategy makes the denominator for this QM a more homogenous group across facility and thus makes the comparison across facility more fair and transparent. Residents are also excluded from the dominator if they are comatose (0.4 percent), have an indwelling catheter (5.8 percent) or ostomy (1.8 percent). These latter exclusions are not considered risk-adjustment exclusions.

Figure 4.11-1
Seasonal (quarterly) variation in QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)



	Q1 2011	Q2 2011	Q3 2011	Q4 2011
Mean	0.4149	0.4175	0.4214	0.4281
Median	0.4146	0.4194	0.4211	0.4265

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 - Quarter 4, 2011 \qm_quarter_1_2 \complete\nh_019_10.log \qm_quarter_2_3\complete\nh_019_10.log \qm_quarter_3_4\complete\nh_019_10.log \qm_quarter_4_5\complete\nh_019_10.log

4.12 QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

4.12.1 Summary of Findings

QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) reports the percentage of long-stay residents whose target assessment indicating the use of indwelling catheters. This QM is risk adjusted based on logistic regression models, and two covariates are used: whether the resident has frequent bowel incontinence, and whether the resident has pressure ulcers at stages 2 through 4 on the prior assessment.

Using MDS 3.0 data, 81.7 percent of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 86.9 percent of the facilities are able to report this QM (i.e., those with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility risk-adjusted QM score is 4.1 percent. The risk-adjusted QM score varies across facility with a standard deviation of 3.3 percent and interquartile range of 4.1 percent.

As reliability checks of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A similar pattern was found for facility rank changes. Most of facilities are with rank changes within three deciles.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures because they reflect similar care processes. We selected QM# 0684 Percent of Residents With a Urinary Tract Infection (Long Stay) and QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the three QMs. Second, we examined the frequency and distribution of missing data. Missing rate is very low for the items used to construct this QM and therefore should not pose a threat to validity of the QM. Last but not the least, we also analyzed the potential geographic (state) variations in this QM. The state of the facility explains about 4.1 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

This QM is risk adjusted based on logistic regression models with two covariates. We examined the impact of partially risk adjustment (adjusting for subsets of the four covariates and on the risk-adjusted QM scores and facility score changes. We also compared coefficients for the covariates across quarter. The results support the validity of the risk-adjustment model. In addition, we explored different model specifications (single-level logistic regression models vs. hierarchical models) for the risk-adjustment model. The results show that changing model specifications has a big impact on facility rank based on the risk-adjusted QM score, as more than half of the facilities with extremely poor performance flagged using the hierarchical model were not flagged using the single-level model.

4.12.2 Background and Introduction to QM

This measure reports the percentage of residents who have had an indwelling catheter in the last 7 days, as noted on the most recent MDS 3.0 assessment, OBRA, PPS or discharge during the selected quarter (3-month period). This measure updates CMS' MDS 2.0 QM on catheter insertions. It is based on data from Minimum Data Set (MDS) 3.0 assessments of long-stay nursing home residents (with cumulative days in the facility greater than 100 days).

A. Background for This QM

A multistate study showed that the prevalence of indwelling urinary catheterization was 12.6 percent at admission to a skilled nursing facility and 4.5 percent at the resident's annual assessment (Rogers, Mody, et al., 2008). Catheters are commonly used for urinary retention, wound management, and in some circumstances, patient comfort. The benefits of limiting catheter use in nursing facilities are well documented in the literature. When not properly maintained and monitored, indwelling catheters can cause chronic pain or infections leading to a greater functional decline and decreased quality of life for the resident (J. K. Gammack, 2003; Quality Measures Management Information System (QMIS), 2002). A thorough assessment of the resident and evaluation of the medical need for the catheter can sometimes decrease or prevent the use of catheters. Indwelling urinary catheterization can frequently causes bacteremia, or in many cases, urinary tract infections, in the elderly. Catheterization causes bacteremia to occur at a rate of 3 to 10 percent of patients per day; a single in and out catheterization may cause bacteremia in as many as 20 percent of patients (Kamel, 2004). The literature suggests that 30 to 40 percent of nursing home-associated infections are related to urinary tract infection and that having an indwelling urinary catheter significantly increases the risk of urinary tract infection (Newman et al., 2004; Rogers, Fries, et al., 2008; Tsan et al., 2008). Many times, residents are admitted to a nursing facility from hospitals with catheters in place, and the facility must make a determination whether or not to continue use of the device.

A thorough assessment of the resident and evaluation of the medical need for the catheter can sometimes decrease or prevent the use of catheters and the risks associated with their use. There are clear benefits to nursing homes conducting a thorough evaluation of the medical need for the catheterization of their residents. A determination regarding continued use or removal should be completed as soon as possible following admission. Nursing facilities need to assess the frequency of urinary catheterization practices to ensure that policies reflect current practice standards, and increase compliance with Centers for Disease Control guidelines for prevention of infection related to catheter use (MedQIC, 2004). The indwelling catheter quality measure can potentially serve as a reminder to facilities of the importance of limiting catheter use (Brega et al., 2007). Overuse of catheters to manage incontinence, other than for short-term periods is a potential sign of suboptimal care and an indication that further assessment and alternative treatment could be offered (Georgiou, Potter, Brocklehurst, Lowe, & Pearson, 2001). Among nursing facility residents, there is evidence that institutional policies and educational programs strongly impact care provider practices.

Studies of catheterization rates have shown there is great variability in this quality measure across states and facilities. Using MDS 2.0 data in the third quarter of 2010, the prevalence of residents with indwelling catheters on the target assessment is 7.0 percent

nationally and ranges from 4.7 percent in Rhode Island to 12.3 percent in Nevada (Centers for Medicare & Medicaid Services, 2012). In a study to measure the quality of urinary continence care in long-term care facilities, catheterization rates were approximately 10 percent in nursing facilities, ranging from 0 to 44 percent among 14 nursing homes where data was collected on the outcome measure (Georgiou et al., 2001). The authors were cautious to note that interpretation of the outcome results required more precise details on case-mix and the definition of outcome measures. In another study looking at state variation in indicators of quality of care in nursing facilities, limited variation among states was observed for urinary catheterization. However, among the risk-adjusted quality scores, the authors observed that urinary catheterization was the quality measure with the largest variance across states, with an approximately twofold difference between the state with the lowest catheterization rate (3.6 percent) to the highest (7.7 percent) (Castle, Degenholtz, & Engberg, 2005).

B. Summary of Differences between MDS 2.0 to MDS 3.0 QM Definitions

Both the MDS 2.0 and MDS 3.0 measures define the numerator as residents with an indwelling catheter The MDS 2.0 capture chronic care residents and the MDS 3.0 captures long-stay residents. In the MDS 2.0, measure residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments and typically have a 30-day average length of stay. In the MDS 3.0 measure, the analogous sample is for long-stay residents, defined as residents with more than 100 cumulative days in facility, and assessments may be discharge, 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments, or admission, quarterly, annual, significant change, or significant correction OBRA assessments (A0310A = 02, 03, 04, 05, 06: OR A0310B = 02, 03, 04, 05: OR A0310F = 10,11).

Residents are counted in the numerator if they are long-stay residents, defined as residents whose length of stay is greater than 100 cumulative days in the facility, and have indwelling catheters (H0100A = 01) on the most recent MDS 3.0 assessment.

The MDS 3.0 measure has several additional exclusion criteria that were not present in the MDS 2.0 measure. While both versions of the measure exclude target assessments that are admission assessments or that have missing data on the item indicating whether an indwelling catheter was present, the MDS 3.0 also excludes the following: target assessments that are readmission/return assessments; assessments that indicate neurogenic bladder (I1550 = [1]) or neurogenic bladder status is missing (I1550 = [-]); and assessments that indicate obstructive uropathy (I1650 = [1]) or obstructive uropathy status is missing (I1650 = [-]). Residents with neurogenic bladder or obstructive uropathy are excluded because these are conditions in which the person is unable to empty the bladder voluntarily or effectively, putting the person at risk or complications, such as overflow incontinence, recurrent infection, vesicoureteral reflux, or autonomic dysflexia.

The MDS 3.0 measure specification applies covariates as risk-adjusters: frequent bowel incontinence on prior assessment (H0400 = 2 or 3), or pressure ulcers at stages 2, 3, or 4 on prior assessment (M0300B1, M0300C1, or M0300D1 is greater than 0). These covariates remain basically unchanged between MDS 2.0 and MDS 3.0, with one exception: MDS 2.0 only used pressure ulcers at stages 3 or 4 as the covariate.

In the national analysis of assessing the reliability of the MDS 3.0 conducted by the RAND Corporation, agreement between MDS 3.0 assessors on bladder and bowel items, including catheter use, was excellent. The average kappa for the gold-standard nurse to gold-standard nurse agreement was 0.949, and the average kappa for the gold-standard nurse to facility nurse agreement was 0.945 (Saliba & Buchanan, 2008).

C. Summary of Analyses

The following sections summarize the basic analyses performed on this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Four general areas are addressed:

- Number of assessments that are included and excluded from the numerator of the QM based on sample restrictions
- Whole sample prevalence of items that are related to the QM
- Findings regarding variability, reportability, reliability, and validity
- Risk-adjustment analyses

4.12.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 1,144,608 (95.1 percent) assessments included in the denominator of QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay), and a total of 59,118 resident episodes (18.3 percent of the total number of long-stay resident episodes) were excluded from the denominator. ***Table 4.12-1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criteria. Note that one resident episode can meet multiple exclusion criteria. The number and percent of resident episodes that met each exclusion criterion are as following:

- a total of 749 (0.06 percent) assessments met exclusion criterion 1: missing data the item indicating presence of indwelling catheter;
- 22,402 (1.9 percent) assessments met exclusion criterion 2: an assessment that was an admission or PPS 5-day or readmission/return assessment;
- 29,954 (2.5 percent) met exclusion criterion 3: neurogenic bladder or missing neurogenic bladder status, and
- 36,666 (3.0 percent) met exclusion criterion 4: obstructive uropathy or missing obstructive uropathy status.

Table 4.12-1
Target assessments used in the calculation of QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

Resident Episodes in the Reporting Period	Frequency[a]	Percentage[a]
Included	1,144,608	95.1%
Excluded—Exclusion 1	1,012	0.1%
Excluded—Exclusion 2	22,402	1.9%
Excluded—Exclusion 3	29,570	2.5%
Excluded—Exclusion 4	36,666	3.0%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Missing data
Exclusion 2 = Wrong target assessment type (admission assessment, PPS 5-day or readmission/return assessment)
Exclusion 3 = Neurogenic bladder or missing data for neurogenic bladder status
Exclusion 4 = Obstructive uropathy or missing data for obstructive uropathy status

[a] Column values may not add up to total since a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident assessments in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log).

Numerator Items – Distribution of Triggering Events

Table 4.12-2 describes the responses to the key item on the MDS 3.0 that contributes to calculation of the numerator for this QM. Item H0100a asks whether an indwelling catheter is used, with a response of "1" indicating yes and a response of "0" indicating no. Across the entire sample of target assessments included in the calculation of this QM, 5.8 percent indicate the presence of an indwelling catheter.

Table 4.12-2
Responses to H0100a (Have Indwelling Catheters) on target assessments not otherwise excluded from analysis

Item	Entry	Frequency	Percentage
H0100a	— = missing	749	0.06%
H0100a	0 = No	1,133,612	94.2%
H0100a	1 = Yes	69,365	5.8%
H0100a	— Total	1,203,726	—

NOTE: H0100a = Indwelling catheter

Analysis date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_020_10_ra.log)

4.12.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with "perfect scores" (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes), which may indicate a "ceiling effect." ***Table 4.12-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0686 is 4.1 percent with a standard deviation of 3.3 percent. This QM has an interquartile range of 4.1 percent. About 13.3 percent of facilities have scores of 0 percent ("perfect" scores).

Table 4.12-3
QM score distribution for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

Risk adjustment used	*n*	Mean score	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
Unadjusted	13,629	4.2%	3.5%	0%	1.9%	3.6%	6.0%	8.8%	13.0%	4.1%
Adjusted	13,527	4.1%	3.3%	0%	1.8%	3.5%	5.9%	8.5%	13.3%	4.1%

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility-level

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_020_10_ra.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examine the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.12-4*** shows the results of this analysis using of MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 86.9 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) using MDS 3.0. This is an improvement over the MDS 2.0 measure, which 85.5 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the new discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011 were discharge assessments, and that 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may improve the reportability of long-stay QMs to some extent. Therefore, the reportability of short-stay QMs may be substantially improved after including the discharge assessment, but the impact on the long-stay QM should be less marked.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay), we report changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.12-4.*** After excluding discharge assessments from the target assessment sample, reportability is slightly improved for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) (86.9 percent with discharge assessments vs. 86.1 percent without discharge assessments).

Table 4.12-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (Percent) of resident episodes in the denominator	1,144,608 (95.1%) [a]	1,123,047 (86.1%) [a]
Number (Percent) of resident episodes in the numerator	47,852 (4.0%) [a]	46,133 (3.8%) [a]
Number (Percent) of facilities able to report this QM	13,629 (86.9%) [b]	13,507 (86.1%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay sample (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one "short-stay" or long-stay residents (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.12-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. From Quarter 1 to Quarter 2, rankings were mostly stable, with nearly 60 percent of facilities maintaining their rank within 10 percentile points, but over 20 percent of facilities shifting by more than 2 deciles. The most recent transition was by far the most stable for facility rank, with 78.6 percent of facilities maintaining the same score and under 7 percent changing by more than 2 deciles.

Table 4.12-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 4.1 percent). Overall, the average change in scores for this measure was a decline of 0.2 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (73.9 percent saw changes of less than one standard deviation), while fewer than 2 percent of facilities saw changes of more than three standard deviations.

Table 4.12-5
Distribution of change in facility rank from one quarter to the next for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	7,506 (58.4%)	7,805 (59.2%)	4,675 (78.6%)
Between 1 and 2 deciles	2,409 (18.7%)	2,373 (18.0%)	883 (14.9%)
Between 2 and 3 deciles	1,477 (11.5%)	1,478 (11.2%)	292 (4.9%)
More than 3 deciles	1,462 (11.3%)	1,520 (11.5%)	98 (1.7%)
Total	12,854	13,176	5,948

NOTES:

Total number of facilities in each column reflects all the facilities that could report the measure in both quarters.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 through Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.12-6
Change in facility score from one quarter to the next for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of Facilities That Declined by More than Three Standard Deviations:	Proportion of Facilities That Declined by between Two and Three Standard Deviations:	Proportion of Facilities That Declined by between One and Two Standard Deviations:	Proportion of Facilities That Declined by Less than One Standard Deviation:	Proportion of Facilities That Improved by Less than One Standard Deviation:	Proportion of Facilities That Improved by between One and Two Standard Deviations:	Proportion of Facilities That Improved by between Two and Three Standard Deviations:	Proportion of Facilities That Improved by More than Three Standard Deviations:
13,289	-0.2%	2.3%	0.68%	2.4%	10.1%	32.1%	41.8%	10.4%	1.9%	0.68%

NOTES:

Number of facilities reporting (13,289) reflects facilities that reported this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCES: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home QMs have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures should be correlated). We examined whether a facility's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures where applicable. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) and the facility's percentile rank on QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) and on QM #0685 Percent of Low Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) (as each of these measures concern processes and outcomes related to the health of the urinary tract) in Quarter 4, 2011. We found that the correlation between the nursing home's percentile rank on QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) and percentile rank on QM #0684 Percent of Residents With a Urinary Tract Infection (Long Stay) and QM #0685 Percent of Low Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) was small ($r = 0.119$) and statistically significant ($p < 0.001$). There was no significant correlation between QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) and QM #0685 Percent of Low Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) ($r = 0.017$, n.s.).[31]

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as state-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether state characteristics might be a source of facility score variation for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) and thus a potential threat to the measure's validity we conducted ANOVA. We also examined interquartile difference between the mean state-level scores across state. The proportion of variance in this measure explained by the state in which facilities are located is 4.0 percent and significant [$F(50, 13{,}476) = 11.2$, $p < .001$].[32] The difference between the mean state-level scores for states at the 25th percentile and the 75th percentile is 1.5 percentage points. Thus, while the majority of the variance in QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) scores is due to factors other than geography, there is a small and significant proportion of the variance that is explained by the respective states in which nursing facilities are located. This feature of the measure warrants future monitoring.

---

[31] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

[32] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\db\quarter_4_5\db160_request)

Missing Data. Missing data represent a potential threat to the validity of a quality measure. Missing data patterns by certain types of residents or assessments indicate that the measure may be not be capturing processes and outcomes for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

In the fourth quarter of 2011, 59,118 residents (4.9 percent) of the long-stay sample were excluded from the construction of this quality measure. After removing those residents with admission assessment or PPS 5-day or readmission/return assessment (these were 22,402 residents: 1.9 percent of the long-stay population), the remaining 67,632 excluded residents (3.0 percent of the long-stay population) were dropped for one or more of the following reasons: neurogenic bladder (29,954 residents), obstructive uropathy (36,666), or missing data on the items used to construct this measure (1,012).

Table 4.12-7 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. There are virtually no missing data on items used to construct this measure: the mean facility-level missing rate is 0.1 percent, and at least 90 percent of facilities have no missing data on these items. Thus, missing data do not represent a potential threat to the validity of this measure.

Table 4.12-7
Distribution of facility-level missing rate for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

n	Mean	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,399	0.09%	1.2%	0%	0%	0%	0%	0%

NOTES:

n = number of facilities who have data for numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request.log)

We further examined the possible relationship between missing data and QM scores. ***Table 4.12-8*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure. In each quartile, the mean facility-level missing rate is approximately one-tenth of 1 percent. Further, there is no correlation between missing data and QM scores (r = -0.0017, *n.s.*).

Table 4.12-8
Distribution of facility-level missing rate for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) by quartile of QM score

Quartile of QM Score	*n*	Mean	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,842	0.09%	1.8%	0%	0%	0%	0%	0%
26%–50%	3,822	0.09%	1.2%	0%	0%	0%	0%	0%
51%–75%	3,883	0.07%	0.7%	0%	0%	0%	0%	0%
76%–100%	3,782	0.09%	1.0%	0%	0%	0%	0%	0%

NOTES:

n = number of facilities who have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2) the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0674 by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.12-9***. For this measure, discharge assessments have virtually no effect on QM score: the means and medians of score change after excluding assessments are both 0 percent.

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. NQF expressed interest in whether seasonal variation might play a role in the score for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay). To address this, we examine the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.12-1***.

Table 4.12-9
Distribution of facility-level score change after excluding discharge assessment for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

n	Mean score change	Std. dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
13,507	0.0%	0.7%	-0.6%	0.0%	0.0%	0.0%	0.3%

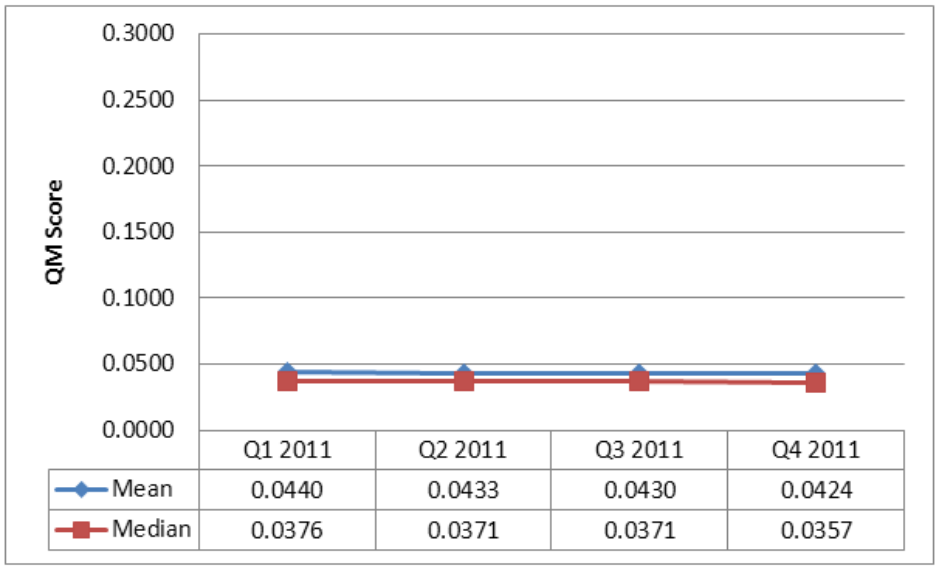
NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Figure 4.12-1
Seasonal (quarterly) variation in QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)



SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 1 - Quarter 4, 2011 \qm_quarter_1_2\complete\nh_019_10.log \qm_quarter_2_3\complete\nh_019_10.log \qm_quarter_3_4\complete\nh_019_10.log \qm_quarter_4_5\complete\nh_019_10.log

The national-level mean for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, was 4.4 percent, 4.4 percent, 4.3 percent, and 4.2 percent. The quarterly national median was 3.8 percent, 3.7 percent, 3.7 percent, and 3.6 percent. While the mean decreased over all four quarters, the median followed the same pattern with the exception that it remained constant from Quarter 2 2011 to Quarter 3 2011. While these early data illustrate no strong seasonal fluctuations, with only 12 months of data available, it is too early to ascertain if the changes in national QM score over time are consistent with seasonal variation, this will need to be confirmed when multiple years of data become available.

4.12.5 Risk Adjustment

QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay) is risk adjusted based on logistic regression models, and two covariates are used: whether the resident has frequent bowel incontinence, and whether the resident has pressure ulcers at stages 2 through 4 on the prior assessment. The frequencies of these two covariates in the long-stay sample are shown in ***Table 4.12-10***. Less than half of the sample (48.7 percent) have frequent bowel incontinence in the prior assessment and a small percent (4.7 percent) of the sample have pressure ulcers at stages 2 through 4. ***Table 4.12-10*** also shows the impact of covariates on whole-sample mean QM score and the number of facilities whose score differs based on the use of covariates. When both covariates— bowel incontinence (covariate 1) or the pressure ulcer covariate (covariate 2)—are included in the risk-adjustment model, the national mean risk-adjusted QM score is 4.2 percent. When only one of the covariates is used, the mean is not affected. However, for facilities with at least a sample size of 30, the use of only the bowel incontinence covariate results in 25.9 percent (n = 3,495) of facilities whose QM score decreases substantially (defined by a difference of 5 percent of the national adjusted mean score) and 31.6 percent (n = 4,269) of facilities whose QM score improves substantially (again, by 5 percent of the national adjusted mean score). The use of only the pressure ulcer covariate results in 19.0 percent (n = 2,559) of facilities whose QM score decreases substantially and 23.2 percent (n = 3,127) of facilities whose QM score increases substantially.

To examine if the performance of the risk-adjustment model for this QM is stable across quarter, we compare the intercepts and coefficients for the covariates for each quarter in 2011. In ***Table 4.12-11***, we show that values for the intercept and for both covariates are very similar across quarter. The model goodness-of-fit statistic—Pseudo R-square—also has comparable values across four quarters. Overall, these results suggest that the performance of the risk adjustment is stable in 2011 and thus support the validity of the risk-adjustment model.

Table 4.12-10
Frequency and impact of covariates used for risk adjustment: QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

Risk-adjustment method	Target assessments with covariate = 1	Mean	SD	Median for adj. QM	# facilities	QM increase of greater than 5% of adjusted mean from using only one covariate	QM decrease of greater than 5% of adjusted mean from using only one covariate
No covariates used	N/A	4.2%	3.5%	3.6%	13,629	4,635 (34.0%)	3,782 (27.7%)
Only covariate 1 used	48.7%	4.3%	3.5%	3.6%	13,494	4,269 (31.6%)	3,495 (25.9%)
Only covariate 2 used	4.7%	4.2%	3.4%	3.6%	13,491	3,127 (23.2%)	2,559 (19.0%)
Covariates 1 and 2 used	—	4.2%	3.4%	3.6%	13,486	N/A	N/A

NOTES:

Covariate 1: Frequent bowel incontinence on prior assessment

Covariate 2: Pressure ulcers at stages 2, 3 or 4 on prior assessment

Total n (denominator) for this QM is 963,959

Facilities with the sample size under 30 are dropped from the facility-level QM analysis

Analysis date: 6/29/12

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db171_request\db171_request.log)

Table 4.12-11
Intercepts and coefficients by quarter of data for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)

Covariate in 2011	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Intercept	-3.653	-3.638	-3.660	-3.677
Bowel incontinence (frequent)	0.521	0.527	0.556	0.554
Pressure ulcers at stages 2, 3, or 4	1.971	1.951	1.970	1.988
Pseudo R-square	0.069	0.068	0.069	0.068

NOTES:

1. Frequent bowel incontinence on prior assessment (H0400 = [2, 3]).

   Covariate = [1] if H0400 = [2, 3]

   Covariate = [0] if H0400 = [0, 1, 9, −].

2. Pressure ulcers at stages 2, 3 or 4 on prior assessment:

   Covariate = [1] if any of the following are true:

   a. M0300B1 = [1, 2, 3, 4, 5, 6, 7, 8, 9], or

   b. M0300C1 = [1, 2, 3, 4, 5, 6, 7, 8, 9], or

   c. M0300D1 = [1, 2, 3, 4, 5, 6, 7, 8, 9]

   Covariate = [0] if M0300B1 = [0, ^] and M0300C1 = [0, ^] and M0300D1 = [0, ^].

   Covariate = missing if M0300B1 = [−] AND M0300C1 = [−] AND M0300D1 = [−].

3. All covariates are missing if no prior assessment is available.

Analysis date:

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (RTI program reference Table 5a 7/19/2011 Memo)

To understand how model specification influences the performance of the risk-adjustment model and the adjusted QM scores, we estimated the risk-adjustment model for this QM using both a single-level model and a hierarchical model. The current specification of the risk-adjustment model for this QM uses single-level logistic regression, which does not take into account the fact that residents are clustered within nursing homes. The hierarchical model, on the other hand, addressed the issue. We estimated the hierarchical model using the specification of multilevel mixed-effects logistic regression. ***Table 4.12-12*** shows that coefficients for the covariates from the single-level model and the ones from the hierarchical model are very similar (0.554 vs. 0.577 for the bowel incontinence covariate; 1.988 vs. 2.022 for the pressure ulcers covariate), but the hierarchical model has a much higher predictive power. The C-statistics indicate that the single-level model has acceptable power for predicting the catheter outcome (C-stat = 0.657) and the hierarchical model has high predictive power for this QM (C-stat = 0.764).

Table 4.12-12
Intercepts and coefficients in Quarter 4, 2011 for QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay): comparison between single-level model and hierarchical generalized linear model

Covariate in Quarter 4, 2011	Single-level model: covariate	Single-level model: 95% confidence interval	Hierarchical model: covariate	Hierarchical model: 95% confidence interval
Bowel incontinence (frequent)	0.554	0.534–0.574	0.577	0.556–0.597
Pressure ulcers at stages 2, 3, or 4	1.988	1.965–2.012	2.022	1.997–2.046
C-statistic	0.657	—	0.764	—

NOTES:

1. Frequent bowel incontinence on prior assessment (H0400 = [2, 3]).

   Covariate = [1] if H0400 = [2, 3]

   Covariate = [0] if H0400 = [0, 1, 9, −].

2. Pressure ulcers at stages 2, 3 or 4 on prior assessment:

   Covariate = [1] if any of the following are true:

   a. M0300B1 = [1, 2, 3, 4, 5, 6, 7, 8, 9], or

   b. M0300C1 = [1, 2, 3, 4, 5, 6, 7, 8, 9], or

   c. M0300D1 = [1, 2, 3, 4, 5, 6, 7, 8, 9]

   Covariate = [0] if M0300B1 = [0, ^] and M0300C1 = [0, ^] and M0300D1 = [0, ^].

   Covariate = missing if M0300B1 = [−] AND M0300C1 = [−] AND M0300D1 = [−].

3. All covariates are missing if no prior assessment is available.

Analysis date: 6/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\\quarter_4_5\db130_request\nh_020_10_xtra_state.log)

The risk-adjusted QM score based on single-level model is different from the ones based on the hierarchical model. For an average facility, the risk-adjusted QM score based on the single-level model is slightly higher than the one based on the hierarchical model (difference = 0.5 percent). The interquartile range for facility-level difference in risk-adjusted QM score between single-level model and hierarchical model is -0.7 to 0.9 percent. Based on these risk-adjusted QM scores, we also compared facility rank, with a particular interest in facilities between the 95th and 100th percentiles (indicating extremely poor quality of care). Among the 444 facilities identified by the single-level model as poor-quality outliners, only 261 (58.8 percent) were identified as poor-quality outliners based on the hierarchical model.[33] These results suggest that the risk-adjustment model for this QM performs differently with different model specifications. As a result, the risk-adjusted QM scores and facility rank based on these models show difference.

[33] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db130_request\db130_request.log)

4.13 QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

4.13.1 Summary of Findings

QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) reports the percentage of long-stay residents whose target assessment indicates daily physical restraints. This subsection presents descriptive analyses and test results for this QM. Almost all (99.9 percent) of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 88.4 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 2.4 percent. The QM score varies across facility with a standard deviation of 4.2 percent and interquartile range of 3.2 percent. The overall low mean QM score and the relatively large variation suggest that this QM can be useful to identify facilities with poor quality of care.

As reliability checks of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A similar pattern was found for facility rank changes. About 90 percent of facilities had rank changes within three deciles.

We conducted several validity tests for this QM. First, we examined the frequency and distribution of missing data. Missing rate is very low for the items used to construct this QM and therefore should not pose a threat to validity of the QM. We also analyzed the potential geographic (State) variations in this QM. The State of the facility explains about 6.9 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

4.13.2 Background and Introduction to QM

The measure reports the percentage of all long-stay residents in nursing facilities during the selected quarter (3-month period) who were physically restrained daily during the 7 days prior to the MDS assessment (which may be OBRA admission, annual, quarterly, significant change, or significant correction; PPS 5-day, 14-day, 30-day, 60-day, 90-day, or readmission/return; or discharge assessment).

A. Summary of QM Background

Restraints can pose serious risks for residents. Restraints are sometimes used to control behavior for people who are disruptive, aggressive, or dangerous, including those with cognitive impairment (Capezuti, Evans, Strumpf, & Maislin, 1996; Castle & Mor, 1998; Sullivan-Marx, Strumpf, Evans, Baumgarten, & Maislin, 1999). The negative outcomes of restraints may include strangulation, loss of muscle tone, decreased bone density (with greater susceptibility for fractures), pressure sores, increased infections, decreased mobility, depression, agitation, loss of dignity, social isolation, incontinence, constipation, functional decline, abnormal changes in body chemistry and muscular function, and in some cases, resident death (Capezuti, Maislin, Strumpf, & Evans, 2002; Castle & Mor, 1998; Centers for Medicare & Medicaid Services, 2003; Evans et al., 1997; Feinsod, Moore, & Levenson, 1997; Parker & Miles, 1997; Sullivan-Marx, 2001; Williams & Finch, 1997). Specifically, the use of trunk restraints is associated with higher

risk for falls and fractures among residents with dementia (Luo, Lin, & Castle, 2011). The use of physical restraints also often constitutes a disproportionate infringement on the autonomy of the resident (Gastmans & Milisen, 2006). The Omnibus Budget Reconciliation Act of 1987 (OBRA 87) specifically grants residents the right to freedom from physical restraints (Wiener et al., 2007). The associated guideline from the Centers for Medicare & Medicaid Services (CMS) states that "The resident has the right to be free from any physical or chemical restraints imposed for the purpose of discipline or convenience and not required to treat the resident's medical symptoms" (Castle, 2002). The use of physical restraints in nursing homes has declined in the past decade (Mor et al., 2011). This decline is seen in the MDS 2.0 data: second quarter 2008 statewide averages for MDS 2.0-based Chronic Care Restraint Quality Measure (QM) range from 1.0 percent in Alaska to 8.9 percent in California, with a 4.3 percent national average (Centers for Medicare & Medicaid Services, 2007), whereas in the third quarter of 2010, the prevalence of residents who were physically restrained was 2.6 percent nationwide, ranging from 0.3 percent in Alaska to 6.1 percent in Louisiana (Centers for Medicare & Medicaid Services, 2012).

The use of restraints also increases the cost of nursing home care. One study examining almost 12,000 residents in 276 facilities in seven States found that higher levels of nursing-assistant time were consistently provided to restrained residents, resulting in increased staff costs to the facilities (Phillips, Hawes, & Fries, 1993). A 1991 report by the Office of the Inspector General (OIG) at CMS found that nursing homes were able to reduce the use of restraints with no increase in cost of care (Kusserow, 1991). Restraints may also impose additional costs on Medicaid; a 2006 analysis of Medicaid reimbursement data for 525 nursing homes found that residents who had experienced greater use of restraints experienced an increased risk of hospitalization (Carter & Porell, 2006).

Despite declining numbers in restraint use, Congress continues to address the issue; the Health Care Fraud Enforcement Act, introduced in 2009, would strengthen the ability of Civil Rights Division of the Department of Justice to investigate unlawful restraint (United States Senate Special Committee on Aging, 2010). In addition to regulations and guidelines, the use of physical restraints in nursing homes may also be reduced by facility-level interventions. Although review of published clinical trials showed insufficient evidence supporting the effectiveness of educational interventions targeting nursing staff for the prevention or reduction of physical restraint use in geriatric long-term care (Mohler, Richter, Kopke, & Meyer, 2011), two clinical trials supported the effectiveness of multicomponent interventions for reducing restraint use without a significant increase in falls, behavioral symptoms, or medication use (Gulpers et al., 2011; Koczy et al., 2011).

B. Summary of Differences between MDS 2.0 and 3.0 QM Definitions

There is only a slight difference between the MDS 2.0 and the MDS 3.0 restraint quality measure. Both measures use section P items in the MDS; the measure is triggered if the code "2" is marked, indicating that the restraint was used daily. The numerator of the MDS 3.0 measure is the number of residents for whom the items in P0100 are marked as (=2), indicating that a resident was restrained daily for subitems b (P0100B—Trunk restraint used in bed), c (P0100C—Limb restraint used in bed), e (P0100 E—Trunk restraint used in chair or out of bed), f (P0100F—limb restraints used in chair or out of bed), or g (P0100G—Chair prevents rising)

during the 7 days prior to the assessment. The denominator includes all residents for whom there is a completed target assessment, only excluding those residents for whom any data from the relevant items on Section P on the MDS 3.0 are missing. There are no other exclusions and the measure is not risk adjusted. The decision not to risk-adjust was because the decision to use physical restraints must be made by medical experts on an individual, case-by-case basis, and there are no broad rules that can be applied.

The MDS 2.0 measure used Section P, Item 4, indicating whether a resident was restrained daily (=2) for subitems c. (Trunk restraint), d. (Limb restraint), or e. (Chair prevents rising) during the 7 days prior to the assessment. Thus, the proposed measure for MDS 3.0 is essentially the same as the current measure from MDS 2.0, except that the MDS 3.0 measure includes additional information about whether the restraint was used while the resident was in bed or if it was used while the resident was out of bed (e.g., in a chair). The refined restraint categories are designed to eliminate confusion about the definition of a restraint and enhance coding accuracy. No changes are anticipated to the facility's performance on the measure due to the small definitional changes, given that the proposed measure contains the same items as the previous measure, only separated out to improve accuracy.

The only relatively major change in the QM definition between MDS 2.0 and MDS 3.0 is related to sample selection for the long-stay population (chronic care in the MDS 2.0 measure). In the MDS 2.0 calculations, residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments (typically capturing residents with stays over 30 days). In the MDS 3.0 calculations, the analogous sample is the long-stay residents, defined as residents with more than 100 cumulative days in facility, and where assessments may be anticipated/not anticipated discharge; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections summarize the basic analyses performed on this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- number of assessments that are included and excluded from the numerator of the QM based on sample restrictions;
- whole sample prevalence of items that are related to the QM; and
- findings regarding variability, reportability, reliability, and validity.

4.13.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 1,202,399 assessments included in the denominator of QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) (approximately 99.9 percent of all assessments); a total of 1,327 resident episodes (0.1 percent of the total number of short stay/long-stay resident episodes) were excluded from the denominator.

Table 4.13-1 shows the proportion of resident episodes excluded from the denominator for this measure's exclusion criterion. For this measure, the only criterion is missing data.

Table 4.13-1
Target assessments used in the calculation of QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	1,202,399	99.9%
Excluded—Exclusion 1	1,327	0.1%
Total number of short-stay resident episodes	1,203,726	__

NOTES:

Exclusion 1 = Missing Data

[a] Column values may not add up to total because a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

Table 4.13-2 shows the frequency of scored items from Section P, Item 1. A score of 2 indicates that the resident was restrained daily during the 7 days prior to the assessment, a score of 1 indicates restraints were used but less than daily, zero indicates that restraints were not used at all, and a dash or a blank indicates that the item was not filled out. The subitems are P0100B—Trunk restraint used in bed (0.1 percent), P0100C—Limb restraint used in bed (0.1 percent), P0100 E—Trunk restraint used in chair or out of bed (1.6 percent), P0100F—limb restraints used in chair or out of bed (0.1 percent), or P0100G—Chair prevents rising (0.9 percent). This measure demonstrates an infrequent use of a dash in completing the restraint items (0.1 percent).

Table 4.13-2
Frequency of items used to calculate QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

Item	Entry	Frequency	Percentage
P0100B	— = Missing	1,240	0.1%
P0100B	0 = Not used	1,201,382	99.8%
P0100B	1 = Used less than daily	157	0.0%
P0100B	2 = Used daily	947	0.1%
P0100B	Total	1,203,726	
P0100C	— = Missing	1,230	0.1%
P0100C	0 = Not used	1,201,256	99.8%
P0100C	1 = Used less than daily	181	0.0%
P0100C	2 = Used daily	1,059	0.1%
P0100C	Total	1,203,726	
P0100E	— = Missing	1,233	0.1%
P0100E	0 = Not used	1,182,618	98.3%
P0100E	1 = Used less than daily	1,653	0.1%
P0100E	2 = Used daily	18,222	1.5%
P0100E	Total	1,203,726	
P0100F	— = Missing	1,244	0.1%
P0100F	0 = Not used	1,201,152	99.8%
P0100F	1 = Used less than daily	226	0.0%
P0100F	2 = Used daily	1,104	0.1%
P0100F	Total	1,203,726	
P0100G	— = Missing	1,254	0.1%
P0100G	0 = Not used	1,191,627	99.0%
P0100G	1 = Used less than daily	1,345	0.1%
P0100G	2 = Used daily	9,500	0.8%
P0100G	Total	1,203,726	

NOTE: P0100B = Bed rail restraint in bed; P0100C = Limb restraint in bed; P0100E = Trunk restraint in chair or out of bed; P0100F = Limb restraint in chair or out of bed; P0100G = Chair prevents rising.

Analysis date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_021_10.log)

4.13.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with "perfect scores" (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes), which may indicate a "ceiling effect." ***Table 4.13-3*** shows the results of this analysis using the QM score Quarter 4, 2011. The mean score for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) is 2.4 percent with a standard deviation of 4.2 percent. This QM has an interquartile range of 3.2 percent. About 50.5 percent of facilities have scores of 0 percent (perfect scores).

Table 4.13-3
QM score distribution for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
13,864	2.4%	4.2%	0%	0%	0%	3.2%	7.1%	50.5%	3.2%

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_021_10.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 long-stay residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examined the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.13-4*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 88.4 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) using MDS 3.0. This is an improvement over the MDS 2.0 measure, which 85.6 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may improve the reportability of long-stay QMs to some extent. Therefore, the reportability of short-stay QMs may be substantially improved after including the discharge assessment, but the impact on the long-stay QM should be less marked.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay), we report changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.13-4***. After excluding discharge assessments from the target assessment sample, reportability is largely unchanged for QM #0687 (88.4 percent with discharge assessments vs. 88.3 percent without discharge assessments).

Table 4.13-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator	1,202,399 (99.9%) [a]	1,201,864 (99.8%) [a]
Number (%) of resident episodes in the numerator	28,909 (2.4%) [a]	29,003 (2.4%) [a]
Number (%) of facilities able to report this QM	13,864 (88.4%) [b]	13,856 (88.3%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay population (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from

quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.13-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. For each pair of quarters in 2011, facility ranks were stable for this measure. In each transition, at least 83.5 percent of facilities maintained the same rank within one decile. Oddly, in each quarter there were many more shifts of more than three deciles than shifts anywhere from 10 percent to 30 percent. Given the proportion of facilities that did not move from quarter to quarter and the stability of the scores themselves (see next analysis), this seems to be a statistical fluke.

Table 4.13-5
Distribution of change in facility rank from one quarter to the next, QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	11,349 (83.5%)	11,705 (85.4%)	11,881 (86,6%)
Between 1 and 2 deciles	612 (4.5%)	495 (3.6%)	436 (3.2%)
Between 2 and 3 deciles	101 (0.7%)	96 (0.7%)	68 (0.5%)
More than 3 deciles	1,524 (11.2%)	1,403 (10.2%)	1,336 (9.7%)
Total	13,586	13,699	13,721

NOTES:

Total *n*'s reflect facilities that meet minimum requirements for public reporting this QM in both quarters.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 through Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.13-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 4.2 percent). Overall, the average change in scores for this measure was a decline of 0.1 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (81.9 percent saw changes of less than one standard deviation), whereas just over 2 percent saw changes of more than three standard deviations.

Table 4.13-6
Change in facility scores from one quarter to the next, QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
13,721	-0.1%	1.7%	1.3%	1.9%	6.3%	14.9%	67.0%	6.1%	1.6%	1.0%

NOTES:

Number of facilities reporting (13,721) reflects facilities that meet minimum requirements for public reporting this QM in both Quarter 3 and Quarter 4, 2011.

Analysis Date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether State characteristics might be a source of facility score variation for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) and thus a potential threat to the measure's validity, we conducted a one-way analysis of variance (ANOVA). We also did a simple examination of the interquartile difference between the mean State-level scores across States. The proportion of variance in this measure explained by the State in which facilities are located is 6.9 percent and significant [$F(51, 13{,}812) = 16.5$, $p < .001$].[34] The difference between the mean State-level scores for States at the 25th percentile and the 75th percentile is 1.4 percentage points. Thus, although the majority of the variance in QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) scores is due to factors other than geography, a small and significant proportion of the variance is explained by the respective States in which nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns indicate that certain types of residents tend to have assessments with missing data in ways that influence the calculation of a quality measure, then that measure may not be capturing processes and outcomes for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

For this measure, resident episodes are excluded only if there are missing data on MDS items regarding restraints (trunk, limb, or chair restraints, indicated by P0100B, P0100C, P0100E, P0100F, and P0100G). In the fourth quarter of 2011, only 1,327 residents (0.1 percent of the long-stay population) were excluded from the construction of this measure owing to missing data.

Table 4.13-7 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. There are virtually no missing data on items used to construct this QM: the mean facility-level missing rate is 0.1 of a percentage point, the median rate is 0 percent, and at least 90 percent of facilities have no missing data on these items. Thus, missing data do not threaten the validity of this measure.

[34] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

Table 4.13-7
Distribution of facility-level missing rate for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,399	0.1%	1.8%	0%	0%	0%	0%	0%

NOTES:

n = number of facilities that have data for the numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request.log)

We further examined the possible relationship between missing data and QM scores. ***Table 4.13-8*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure. Please note that because of the low prevalence of this measure, the first and second quartiles have been combined. The mean facility-level missing-data rate is near-zero for each of the three groups examined in this analysis. In the highest quartile, the mean missing-data rate is slightly higher (0.2 percent), which may account for the weak and significant correlation between missing data and QM scores for this measure ($r = 0.1419$, $p < .001$). Again, this relationship is likely an artifact of the low prevalence and narrow variance of both the missing-data rates and the scores for this measure.

Table 4.13-8
Distribution of facility-level missing rate for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) by quartile of QM score

Quartile of QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–50%[a]	8,263	0.1%	0.9%	0%	0%	0%	0%	0%
51%–75%	3,294	0.1%	0.8%	0%	0%	0%	0%	0%
76%–100%	3,841	0.2%	2.7%	0%	0%	0%	0%	0%

NOTES:

[a] Facilities cannot be further divided into quartiles due to high prevalence of 0.

Total *n* = number of facilities that have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2) the

completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.13-9.*** The mean nursing home–level QM score and distributions remained largely unchanged after excluding discharge assessments from the set of eligible target assessments. Thus, the use of the discharge assessment in the QM score calculations does not meaningfully alter the distribution of scores among facilities.

Table 4.13-9
Distribution of facility-level score change after excluding discharge assessment for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
13,856	0.0%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

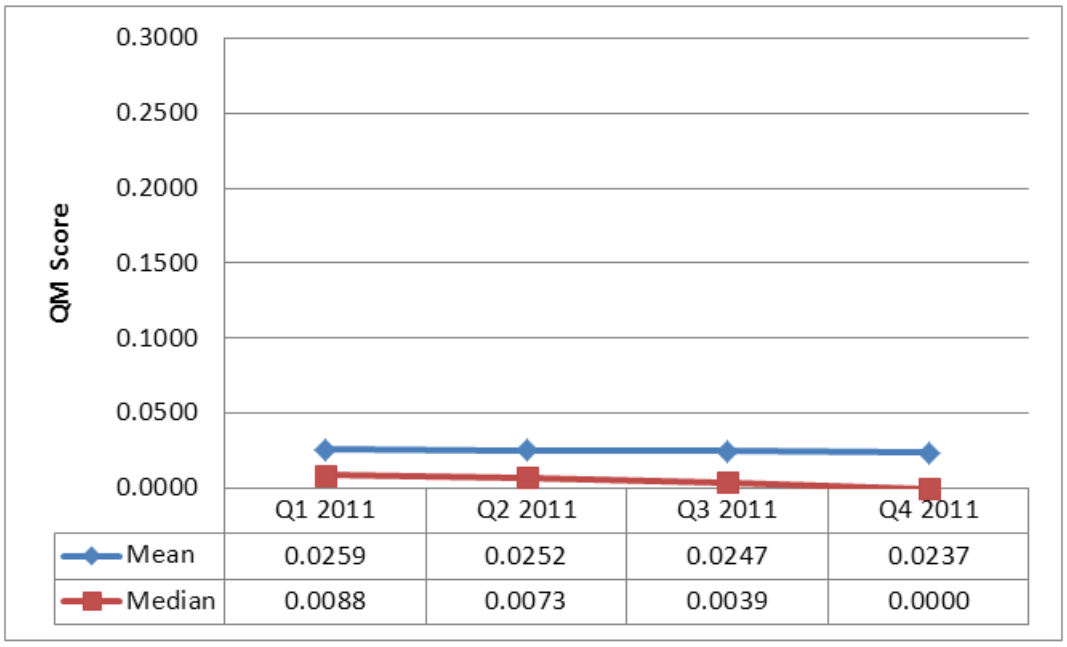
SOURCE: RTI analysis of MDS 3.0 data for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address this interest in whether seasonal variation might play a role in the score for NQF #0687 Percent of Residents Who Were Physically Restrained (Long Stay), we examined the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.13-1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, was 2.6 percent, 2.5 percent, 2.5 percent, and 2.4 percent. The quarterly national median was 0.9 percent, 0.7 percent, 0.4 percent, and 0.0 percent. Both mean and median decreased across all four quarters. With only 12 months of data available, it is too early to ascertain whether the changes in national QM score over time are consistent with seasonal variation; this will need to be confirmed when multiple years of data become available.

Figure 4.13-1
Seasonal (quarterly) variation in QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay)



SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 1 - Quarter 4, 2011 \qm_quarter_1_2\complete\nh_021_10.log \qm_quarter_2_3\complete\nh_021_10.log \qm_quarter_3_4\complete\nh_021_10.log \qm_quarter_4_5\complete\nh_021_10.log

4.13.5 Risk Adjustment

QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay) was endorsed by NQF without denominator exclusion and model-based risk adjustment. The choice to use physical restraints on a resident should be made by a medical professional in concert with the family on a case-by-case basis. There are no broad policy rules that can be applied to risk adjust whether or not a resident is restrained.

4.14 QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

4.14.1 Summary of Findings

QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) reports the percentage of residents whose need for help with late-loss Activities of Daily Living (ADLs) has increased. This subsection presents descriptive analyses and test results for this QM. We found that 81.7 percent of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 81.2 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score was 16.7 percent. The QM score varied across facilities, with a standard deviation of 9.3 percent and interquartile range of 11.8 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

As reliability checks of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation, with a very low proportion of changes being greater than three standard deviations. A slightly different pattern was found for facility rank changes. From quarter to quarter in 2011, about 20 percent of facilities had rank changes of more than three deciles.

We conducted several validity tests for this QM. First, we analyzed the correlations between QMs within a "care process group." The hypothesis was that changes in one QM within a measure group should be correlated with changes in other measures because they reflect similar care processes. We selected QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) to conduct this correlation analysis. Findings from this analysis showed significant but weak correlations among the two QMs (correlation $r = 0.111$). Second, we examined the frequency and distribution of missing data. Missing rate is about 1 percent for the items used to construct this QM and therefore should not pose a threat to validity of the QM. Last but not least, we also analyzed the potential geographic (State) variations in this QM. The State of the facility explains about 6 percent of the variation in this QM. The QM score only slightly decreased from quarter to quarter in 2011.

4.14.2 Background and Introduction to QM

This measure reports the percentage of long-stay residents whose need for help with late-loss ADLs has increased when compared with the prior assessment.

A. Summary of Background on This QM

Increasing ADL dependence is associated with decreased quality of life in older adults, and has been shown to be a risk factor for complications, including pressure ulcers and hospitalizations (Colorado Foundation for Medical Care, 2007). Some ADL loss can result from a resident's clinical conditions (e.g., poor nutrition, prolonged bed rest, injuries, severe cognitive impairment) (Degenholtz, Rosen, Castle, Mittal, & Liu, 2008). ADL deterioration can also result from inadequate nursing care and failure to initiate appropriate rehabilitative or supportive care to mitigate the loss of ADLs. The associations between ADL decline and nurse turnover and nursing home characteristics and location have been established in the literature (Horn et al., 2010; Phillips, Shen, Chen, & Sherman, 2007). In addition, one study found that effects of nursing homes on ADLs tended to be weaker for residents with low cognitive function (Wang et al., 2010).

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 ADL ability 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. Although 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 found evidence that the decline shown using MDS data reflects the natural course of decline (Chen, Chan, Kiely, Morris, & Mitchell, 2007; McConnell, Pieper, Sloane, & Branch, 2002). However, some research suggests that case mix factors such as

level of cognitive impairment, medical factors, and race/socioeconomic status may also contribute to the extent of decline observed in a facility (Ang, Au, Yap, & Ee, 2006; Chen et al., 2007; McConnell et al., 2002; Peek, Coward, Henretta, Duncan, & Dougherty, 1997; Phillips et al., 2007). For some patients, research indicates that functional decline may be lessened by specialized nursing care or other interventions (Boltz, Resnick, Capezuti, Shuluk, & Secic, 2012).

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) (Bellows & Halpin, 2008), 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 (Abt Associates Inc. et al., 2007). To address this, the Rand Corporation developed and tested new ADL measures for inclusion in the MDS, under contract to CMS (Saliba & Buchanan, 2008). However, to avoid undue burden on 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.

This measure is intended to prompt facilities to evaluate whether their long-stay residents are experiencing avoidable ADL declines and, if so, develop approaches to help residents improve or maintain their functional abilities.

B. Summary of Differences between MDS 2.0 and MDS 3.0 QM Definitions

QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) reports the percentage of long-stay residents whose need for help with late-loss ADLs, as reported in the target quarter's assessment, has increased when compared with the previous assessment. 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. The four late-loss ADLs are bed mobility, transfers, eating, and toileting. The measure is calculated by comparing the change in each item between the target MDS assessment (OBRA, PPS, or discharge) and a previous assessment (OBRA, PPS, or discharge). A decline in late-loss ADLs is defined as an increase by two or more points in one late-loss ADL, or increase in one point in two or more late-lost ADLs based on the ADL data available. The numerator is the number of long-stay residents who have an MDS assessment (OBRA, PPS, or discharge) reporting a defined amount of decline when compared with a previous assessment (OBRA, PPS, or discharge). The denominator includes all long-stay residents with a selected target and prior assessment during the quarter and who did not meet the exclusion criteria.

Excluded from the measure are residents that are comatose or totally dependent on all late-loss ADLs, or have a prognosis of life expectancy less than 6 months, or are enrolled in hospice care on the target assessment, or do not meet the criteria for decline in late-loss ADLs, or there are missing data on any of the four late-loss ADL items.

The MDS 2.0 QM has been demonstrated to be problematic by previous research. Concerns over the QM include evidence that the measure has been shown to be heavily influenced by seasonal changes in resident case mix, and that it also is sensitive to State Medicaid policies and payment formulae (i.e., higher levels of ADL decline are reported in States that incorporate ADL decline into their payment formulae) (Ang et al., 2006; Bellows & Halpin, 2008; Chen et al., 2007; McConnell et al., 2002; Phillips et al., 2007). RAND recommended changes to the items to improve reliability (Saliba & Buchanan, 2008). However, because the QM is used in RUGs, changes to the MDS 3.0 were limited to clarifying instructions for scoring the items, and the addition of a response category, "Activity occurred only once or twice."

Note that the overall sample definitions also changed from MDS 2.0 to MDS 3.0. For the MDS 2.0, residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments. For the MDS 3.0, the analogous sample is for long-stay residents, defined as residents with more than 100 cumulative days in a facility, and assessments may be discharge; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections summarize the basic analyses performed on this QM using MDS 3.0 data for the QM for Quarter 4, 2011 (unless otherwise specified). Three general areas are addressed:

- number of assessments that are included and excluded from the numerator of the QM based on sample restrictions;
- whole sample prevalence of items that are related to the QM; and
- findings regarding variability, reportability, reliability, and validity.

4.14.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 967,915 (81.7 percent) assessments included in the denominator of QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay), and a total of 235,811 resident episodes (19.3 percent of the total number of long-stay resident episodes) were excluded from the denominator. ***Table 4.14-1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criteria. Note that one resident episode can meet multiple exclusion criteria. The number and percentage of resident episodes that met each exclusion criterion are as follows:

- A total of 16,615 (1.4 percent) assessments met exclusion criterion 1: missing data on items used to construct the numerator;
- 202,391 (17.1 percent) assessments met exclusion criterion 2: residents showed high level of ADL dependence on prior assessment;
- 4,471 (0.3 percent) assessments met exclusion criterion 3: comatose state;

- 47,267 (4.0 percent) assessments met exclusion criterion 4: a prognosis of life expectancy less than 6 months on the target assessment; and
- 74,782 (6.3 percent) assessments met exclusion criterion 5: hospice care status.

Table 4.14-1
Target assessments used in the calculation of QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	967,915	81.7%
Excluded—Exclusion 1	16,615	1.4%
Excluded—Exclusion 2	202,391	17.1%
Excluded—Exclusion 3	4,471	0.3%
Excluded—Exclusion 4	47,267	4.0%
Excluded—Exclusion 5	74,782	6.3%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Missing Data

Exclusion 2 = Residents are identified as highly dependent on prior assessment.

Exclusion 3 = Comatose

Exclusion 4 = Prognosis of life expectancy is less than 6 months on target assessment.

Exclusion 5 = Hospice care

[a] Column values may not add up to total because a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items—Distribution of Triggering Events

Table 4.14-2 shows the frequency and percentage of triggering events (items that can trigger the numerator of the QM). The triggering items are G0110A1 (bed mobility), G0110B1 (transfer), G0110H1 (eating), G0110I1 (toileting). Response categories for each of these items ranges from a code of 0 (independent) to a code of 4 (total dependence). Code 7 is used if an activity occurred only once or twice; code 8 indicates the activity did not occur during the look-back period. Overall, the triggering events were distributed across all scores, and do not show a floor or ceiling effect. The percentage of residents that are coded as being independent on these late-loss ADL items ranges from 10.2 percent (toileting) to more than 30 percent (eating). For three of the late-loss ADLs, the highest frequency code for each of the items is a code 2 or 3, indicating that limited or extensive assistance is required. Roughly 60 percent of residents score a 2 or 3 for Item G0110A1 (bed mobility), 55 percent score 2 or 3 for Item G0110B1 (transfer), and 56.5 percent score 2 or 3 for Item G0110I1 (toileting). The exception is Item G0110H1 (eating), on which the majority of patients score 1, indicating that supervision is required.

Table 4.14-2
Frequency and percentage of responses on items important for the numerator of QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

Item	Entry	Frequency	Percentage	Cumulative %
G0110A1	— = Missing	2,141	0.2%	0.2%
G0110A1	0 = Independent – no help or staff oversight at any time	192,014	16.2%	16.4%
G0110A1	1 = Supervision – oversight, encouragement, or cueing	83,513	7.1%	23.4%
G0110A1	2 = Limited assistance – resident highly involved in activity; staff provide guided maneuvering	141,373	11.9%	35.4%
G0110A1	3 = Extensive assistance – resident involved in activity; staff provide weight-bearing support	572,849	48.4%	83.7%
G0110A1	4 = Total dependence – full staff performance every time during entire 7-day period	191,786	16.2%	99.9%
G0110A1	7 = Activity occurred only once or twice – activity did occur but only once or twice	559	0.1%	100.0%
G0110A1	8 = Activity did not occur – activity (or any part of the ADL) was not performed by resident or staff at all over the entire 7-day period	410	0.0%	100.0%
G0110A1	Total	1,184,645	—	—
G0110B1	— = Missing	2,120	0.2%	0.2%
G0110B1	0 = Independent – no help or staff oversight at any time	154,773	13.1%	13.2%
G0110B1	1 = Supervision – oversight, encouragement, or cueing	91,004	7.7%	20.9%
G0110B1	2 = Limited assistance – resident highly involved in activity; staff provide guided maneuvering	148,532	12.5%	33.5%
G0110B1	3 = Extensive assistance – resident involved in activity; staff provide weight-bearing support	503,218	42.5%	75.9%
G0110B1	4 = Total dependence – full staff performance every time during entire 7-day period	260,126	22.0%	97.9%
G0110B1	7 = Activity occurred only once or twice – activity did occur but only once or twice	9,061	0.8%	98.7%
G0110B1	8 = Activity did not occur – activity (or any part of the ADL) was not performed by resident or staff at all over the entire 7-day period	15,811	1.3%	100.0%
G0110B1	Total	1,184,645	—	—

(continued)

Table 4.14-2 (continued)
Frequency and percentage of responses on items important for the numerator of QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

Item	Entry	Frequency	Percentage	Cumulative %
G0110H1	— = Missing	2,170	0.2%	0.2%
G0110H1	0 = Independent – no help or staff oversight at any time	359,600	30.4%	30.5%
G0110H1	1 = Supervision – oversight, encouragement, or cueing	361,555	30.5%	61.1%
G0110H1	2 = Limited assistance – resident highly involved in activity; staff provide guided maneuvering	123,033	10.4%	71.4%
G0110H1	3 = Extensive assistance – resident involved in activity; staff provide weight-bearing support	172,455	14.6%	86.0%
G0110H1	4 = Total dependence – full staff performance every time during entire 7-day period	164,251	13.9%	99.9%
G0110H1	7 = Activity occurred only once or twice – activity did occur but only once or twice	676	0.1%	99.9%
G0110H1	8 = Activity did not occur – activity (or any part of the ADL) was not performed by resident or staff at all over the entire 7-day period	905	0.1%	100.0%
G0110H1	Total	1,184,645	—	—
G0110I1	— = Missing	2,178	0.2%	0.2%
G0110I1	0 = Independent – no help or staff oversight at any time	120,753	10.2%	10.4%
G0110I1	1 = Supervision – oversight, encouragement, or cueing	72,152	6.1%	16.5%
G0110I1	2 = Limited assistance – resident highly involved in activity; staff provide guided maneuvering	124,672	10.5%	27.0%
G0110I1	3 = Extensive assistance – resident involved in activity; staff provide weight-bearing support	545,276	46.0%	73.0%
G0110I1	4 = Total dependence – full staff performance every time during entire 7-day period	314,864	26.6%	99.6%
G0110I1	7 = Activity occurred only once or twice – activity did occur but only once or twice	699	0.1%	99.7%
G0110I1	8 = Activity did not occur – activity (or any part of the ADL) was not performed by resident or staff at all over the entire 7-day period	4,051	0.3%	100.0%
G0110I1	Total	1,184,645	—	—

NOTE: G0110A1 = Bed mobility self-performance; G0110B1 = Transfer self-performance; G0110H1 = Eating self-performance; G0110I1 = Toilet use self-performance.

Analysis date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_022_10.log)

4.14.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with perfect scores (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes) which may indicate a ceiling effect. ***Table 4.14-3*** shows the results of this analysis using the QM score for Quarter 4, 2011, QM values. The mean score for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) is 16.7 percent with a standard deviation of 9.3 percent. This QM has an interquartile range of 11.8 percent. About 1.0 percent of facilities have scores of 0 percent (perfect scores).

Table 4.14-3
QM score distribution for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
12,735	16.7%	9.3%	6.3%	10%	15.4%	21.8%	29.0%	1.0%	11.8%

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_022_10.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 long-stay residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examine the percentage of nursing homes that can report each measure (referred to as the QM "reportability"). ***Table 4.14-4*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 81.2 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) using MDS 3.0. This is comparable to the MDS 2.0 measure, which 81.8 percent of facilities were able to report in Quarter 1, 2006.[35]

[35] SOURCE: MDS 3.0, RTI analysis of Quarter 4 2010 and Quarter 1, 2011 data (qm_summary_q1_q2.xlsx); MDS 2.0, University of Colorado Quarter 1 2006 data

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may improve the reportability of long-stay QMs to some extent. Therefore, the reportability of short-stay QMs may be substantially improved after including the discharge assessment, but the impact on the long-stay QM should be less marked.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay), we report changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.14-4***. After excluding discharge assessments from the target assessment sample, reportability is largely unchanged for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) (81.2 percent with discharge assessments vs. 81.1 percent without discharge assessments).

Table 4.14-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Have Increased (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator	967,915 (80.4%) [a]	970,065 (80.6%) [a]
Number (%) of resident episodes in the numerator	160,052 (13.3%) [a]	160,115 (13.3%) [a]
Number (%) of facilities able to report this QM	12,735 (81.2%) [b]	12,714 (81.1%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay population (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db135_request\db135_request.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance.

We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.14-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. Facility rankings on this measure were relatively unstable. For each quarter, fewer than half of the facilities maintained a rank within the same decile in the next quarter. Each quarter also saw about one-fifth of facilities moving more than three deciles in rank. This relative instability is difficult to explain—particularly in light of subsequent analyses (see below) that show that measure scores did not change by a large amount from Quarter 3 to Quarter 4 of 2011. The variability of this measure and its effect on facility ranking should be monitored in the future.

Table 4.14-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 9.3 percent). Overall, the average change in scores for this measure increased by 0.2 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (74 percent saw changes of less than one standard deviation), whereas just over 1 percent saw changes of more than three standard deviations.

Table 4.14-5
Distribution of change in facility rank from one quarter to the next, QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	5,628 (47.2%)	5,800 (47.4%)	5,817 (47.0%)
Between 1 and 2 deciles	2,389 (20.0%)	2,361 (19.3%)	2,429 (19.6%)
Between 2 and 3 deciles	1,602 (13.4%)	1,643 (13.4%)	1,618 (13.1%)
More than 3 deciles	2,312 (19.4%)	2,443 (20.0%)	2,521 (20.4%)
Total	11,931	12,247	12,385

NOTES:

Total *n*'s reflect facilities that meet minimum requirements for public reporting this QM in both quarters.

Columns do not necessarily total 100.0% because of rounding.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 through Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.14-6
Change in facility score from one quarter to the next, QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
12,385	0.2%	9.1%	0.5%	2.0%	10.3%	38.3%	36.5%	9.7%	2.0%	0.8%

NOTES:

Number of facilities reporting (12,385) reflects facilities that meet minimum requirements for public reporting this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures, should be correlated). Following this reasoning, facilities should perform similarly on quality measures that reflect the quality of similar care processes (i.e., their percentile ranking on any of these measures should be correlated). We examined whether a nursing home's percentile rank on one quality measure in a measure group was correlated with its percentile rank on another quality measure in the same clinically related group. The analyses are based on facilities' risk-adjusted measures where applicable. Among nursing homes that could report both related measures, we calculated the correlation between the facility's percentile rank on QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) and the facility's percentile rank on QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) in Quarter 4, 2011, given that both of these measures are concerned with the physical abilities of residents. We found that the correlation between the nursing home's percentile rank on QM #0676 Percent of Residents Who Self-Report Moderate to Severe Pain (Short Stay) and percentile rank on QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) was small ($r = .148$) and statistically significant ($p < 0.001$). [36]

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether State characteristics might be a source of facility score variation for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) and thus a potential threat to the measure's validity, we conducted a one-way analysis of variance (ANOVA). We also did a simple examination of the interquartile difference between the mean State-level scores across States. The proportion of variance in this measure explained by the State in which facilities are located is 6.0 percent and significant [$F(50, 12{,}684) = 18.7$, $p < .001$].[37] The difference between the mean State-level scores for States at the 25th percentile and the 75th percentile is 3.3 percentage points. Thus, although the majority of the variance in QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) scores is due to factors other than geography, a small and significant proportion of the variance is explained by the respective States in which nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns indicate that certain types of residents tend to have assessments with

---

[36] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db173_request\db173_request.log)

[37] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

missing data in ways that influence the calculation of a quality measure, then that measure may not be capturing processes and outcomes for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

In Quarter 4 of 2011, there were 235,811 residents (19.6% of the long-stay population) who were excluded from the construction of this measure. Of these, 16,615 (1.4% of the population) were excluded because of missing data on items used to construct the measure. The remaining resident exclusions were due to presence of one or more of the following conditions: high ADL dependence (202,391), coma (4,471), prognosis of less than 6 months to live (47,267), and/or hospice care (74.782).

Table 4.14-7 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. Missing data on the items used to calculate this measure are rare. The mean facility-level missing rate is 1.5 percent, but median and the 75th percentile are both 0 percent, suggesting that relatively few facilities are responsible for the handful of missing data on these items. It appears that missing data are too rare to present a threat to the validity of this measure.

Table 4.14-7
Distribution of facility-level missing rate for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,290	1.5%	8.2%	0%	0%	0%	0%	1.8%

NOTES:

n = number of facilities that have data for the numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request.log)

We further examined the possible relationship between missing data and QM scores. ***Table 4.14-8*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure. There is a weak, positive correlation between facility-level missing-data rates and QM scores for this measure ($r = 0.0218$, $p = 0.007$). Stratifying mean missing-data rates by quartile, we see that this relationship might not be linear (the first and the fourth quartiles have the highest mean rate). It is also possible that the relationship is driven by the slightly higher mean rate and wider variance in missing-data in the fourth quartile of scores. However, missing data are still relatively rare for items used to construct this measure and account for less than one-tenth of 1 percent of the variance in QM scores for this measure, so it is likely not an important concern.

Table 4.14-8
Distribution of facility-level missing rate for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) by quartile of QM score

Quartile of QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,823	1.7%	0.0%	0%	0%	0%	0%	1.5%
26%–50%	3,822	1.0%	5.9%	0%	0%	0%	0%	1.4%
51%–75%	3,802	1.1%	6.5%	0%	0%	0%	0%	1.6%
76%–100%	3,806	1.8%	8.7%	0%	0%	0%	0%	2.6%

NOTES:

Total n = number of facilities that have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2) the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.14-9.*** The mean nursing home–level QM score and distributions change slightly before and after excluding discharge assessments from the set of eligible target assessments. The average facility score changes by less than one-tenth of 1 percent (with a median change of 0 percent) when discharge assessments are excluded. Thus, the use of the discharge assessment in the QM score calculations does not meaningfully alter the distribution of scores among facilities.

Table 4.14-9
Distribution of facility-level score change after excluding discharge assessment for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
12,714	0.0%	1.7%	-2.1%	-0.8%	0.0%	0.6%	2.0%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

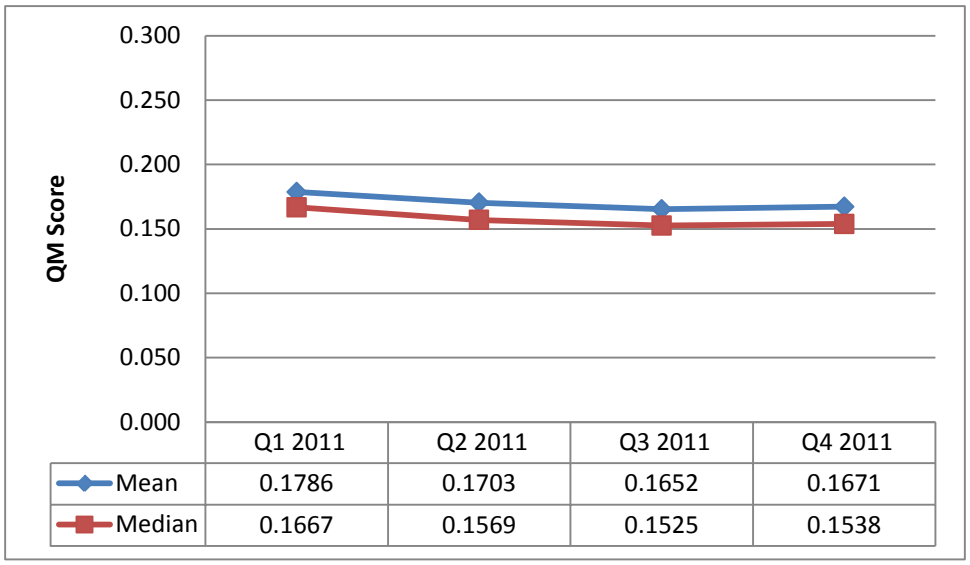
SOURCE: RTI analysis of MDS 3.0 data for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address this interest in whether seasonal variation might play a role in the score for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay), we examined the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4.14-1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011 and Quarter 4, 2011, was 17.9 percent, 17.0 percent, 16.5 percent, and 16.7 percent. The quarterly national median was 16.7 percent, 15.7 percent, 15.3 percent, and 15.4 percent. Both mean and median decreased from Quarter 1, 2011, to Quarter 3, 2011, and increased slightly from Quarter 3, 2011, to Quarter 4, 2011. With only 12 months of data available, it is too early to ascertain whether the changes in national QM score over time are consistent with seasonal variation; this will need to be confirmed when multiple years of data become available.

Figure 4.14-1
Pattern comparison before and after implementing MDS 3.0 for QM #0688 Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay)



	Q1 2011	Q2 2011	Q3 2011	Q4 2011
Mean	0.1786	0.1703	0.1652	0.1671
Median	0.1667	0.1569	0.1525	0.1538

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 - Quarter 4, 2011 \qm_quarter_1_2\complete\nh_022_10.log \qm_quarter_2_3\complete\nh_022_10.log \qm_quarter_3_4\complete\nh_022_10.log \qm_quarter_4_5\complete\nh_022_10.log

4.14.5 Risk Adjustment

QM #0688, Percent of Residents Whose Need for Help With Activities of Daily Living Has Increased (Long Stay) is not risk adjusted. Although previous research has indicated that the measure might benefit from risk adjustment, results from testing of risk-adjustment models using the MDS 2.0 measure specifications were poor (R-square=0.005). This is less than the 0.01 standard established in the analyses (Brega et al., 2007). 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 (Arling, Lewis, Kane, Mueller, & Flood, 2007). 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 0.12 (Phillips et al., 2007). However, the above analyses addressing public comments voicing concerns about differences in risk associated with different patient characteristics and combinations of diagnoses suggest potential adjusters that should be tested.

4.15 QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)

4.15.1 Summary of Findings

QM #0689, Percent of Residents Who Lose Too Much Weight (Long Stay) reports the percentage of residents who had a weight loss of 5 percent or more in the last month or 10 percent or more in the last two quarters who were not on a physician prescribed weight-loss regimen. This subsection presents descriptive analyses and test results for this QM. We found that almost all (97.4 percent) of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 87.7 percent of the facilities are able to report this QM (i.e. with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 7.1 percent. The QM score varied across facilities, with a standard deviation of 4.6 percent and interquartile range of 5.8 percent.

As reliability checks of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A slightly different pattern was found for facility rank changes. From quarter to quarter in 2011, about 20 percent of facilities had rank changes of more than three deciles.

We conducted several validity tests for this QM. First, we examined the frequency and distribution of missing data. Missing rate is less than 1 percent for the items used to construct this QM and therefore should not pose a threat to validity of the QM. We also analyzed the potential geographic (State) variations in this QM. The State of the facility explains about 3.3 percent of the variation in this QM. The QM score was largely stable from quarter to quarter in 2011.

4.15.2 Background and Introduction to QM

This measure updates CMS's MDS 2.0 QM on patients who lose too much weight. This measure captures the percentage of long-stay residents who had a weight loss of 5 percent or more in the last month or 10 percent or more in the last 6 months who were not on a physician-prescribed weight-loss regimen noted on an MDS assessment (OBRA, PPS, or discharge) during the selected quarter (3-month period).

To address seasonal variation, the proposed measure uses a two-quarter average for the facility. Long-stay residents are those who have been in nursing care for more than 100 days. The measure is restricted to this population, which has long-term care needs, rather than the short-stay population who are discharged within 100 days of admission.

A. Summary of QM Background

Nursing home residents often have chronic diseases and functional impairments that impair proper nutrition and hydration (Morley, 2007); (Bourdel-Marchasson, 2010; Sloane, Ivey, Helton, Barrick, & Cerna, 2008) and require interventions by facility staff (Morley, 2007). Elderly individuals with weight loss are at higher risk for functional decline, hip fracture (Ensrud et al., 2003; Langlois, Harris, Looker, & Madans, 1996; Langlois et al., 2001) and mortality (Amador, Al Snih, Markides, & Goodwin, 2006; Covinsky et al., 1999; Keller & Ostbye, 2005; Kiely & Flacker, 2000; Ryan, Bryant, Eleazer, Rhodes, & Guest, 1995; Sullivan et al., 2002; Wedick, Barrett-Connor, Knoke, & Wingard, 2002). Consequences of weight loss also may include muscle wasting, infections, and increased risk of pressure ulcers. Detecting and preventing weight loss is central to ensuring appropriate nutritional intake.

Prevalence estimates of poor nutrition and unintentional weight loss among people in institutions vary from 2 percent to 41 percent (Pauly, Stehle, & Volkert, 2007); dehydration is also common (Amella, 2004). Using MDS 2.0 data for July to September 2010, the national prevalence of too much weight loss in nursing homes was 8.6 percent, ranging from a low average of 6.3 percent in Iowa to a high average of 11.2 percent in West Virginia (Centers for Medicare & Medicaid Services). The national percentage associated with too much weight loss fluctuated somewhat between 2003 and 2009, with a modest downward pattern (American Health Care Association, 2011). Preliminary testing of QM using MDS 3.0 data and a sample of nursing homes estimated a prevalence that was virtually the same as the prevalence estimate based on the MDS 2.0 QM (8.3 percent vs. 8.0 percent) (Saliba & Buchanan, 2008).

Various chronic illnesses are associated with malnutrition, including cancer, diabetes, depression, and COPD (Huffman, 2002). Medications, oral health problems (such as missing teeth), dysphagia, and dementia can complicate nutrition and hydration. Between 40 percent and 60 percent of nursing home residents have swallowing disorders, often related to dementia (Kayser-Jones & Pengilly, 1999). Medications may cause nausea, anxiety, constipation, and lack of appetite. Depression has been identified as the "most common reversible illness" associated with malnutrition (Sloane et al., 2008). Dehydration is a major factor in weight loss in an estimated 10 percent of nursing home residents (Feinsod et al., 2004; Kaldy, 2000; Smith, 2006). The Council for Nutritional Clinical Strategies in Long-Term Care, an expert panel of interdisciplinary thought leaders representing academia and the medical community, derived a structured approach aimed at improving management of malnutrition in long-term care settings, using literature review and consensus development. The Clinical Guide to Prevent and Manage Malnutrition in Long-Term Care is based on a best-evidence approach to the management of nutritional problems in long-term care. The recommendations were determined by consensus process by the Council for Nutritional Clinical Strategies in Long-Term Care, and clinical triggers were reviewed by an independent GSA peer-review committee. The parameters for identifying malnutrition in nursing homes were derived from OBRA 1987 guidelines, including

involuntary weight loss of greater than 5 percent in 30 days or 10 percent in 180 days, which is used as the trigger in this quality measure (Thomas, Ashmen, Morley, & Evans, 2000).

A loss of 5 percent or more of body weight in 1 month or 10 percent or more over 6 months is usually considered unhealthy (Thomas et al., 2000). Too much weight loss can make a person weak, change how medicine works in the body, or cause the skin to break down, which can lead to pressure ulcers. Too much weight loss may mean that the resident is ill, refuses to eat, is depressed, or has a medical problem that makes eating difficult (e.g., weakness caused by a stroke). It could also mean that the resident is not being fed properly; that his or her medical care is not being properly managed; or that the nursing homes' nutrition program is poor. To help prevent unhealthy weight loss, it is important that the resident's diet be balanced and nutritious, and that staff spend enough time feeding residents who cannot feed themselves.

With the increase in obesity, it may be necessary for some residents to lose weight for medical reasons. The rate of obesity in nursing homes increased in the past decade to more than 25 percent (Lapane & Resnik, 2005). In these cases, the medical staff may plan in advance for the resident to lose weight on a special weight loss program, but the person should not lose more than 5 percent of body weight in 1 month. The current MDS weight-loss quality indicator was found to be reliable in differentiating nursing homes with a lower prevalence of weight loss from those with a higher prevalence. Significantly more residents were at risk for weight loss in high-weight-loss nursing homes, according to multiple measures, most notably low oral intake as measured by the MDS and direct observations by research staff. One care process that consistently differentiated care in low-weight-loss nursing homes from that in high-weight-loss nursing homes across all-risk group comparisons was the presence of verbal prompting or social interaction during meals. Specifically, staff in low-weight-loss nursing homes provided verbal prompting and social interaction during meals to a significantly greater proportion of all participants and, in particular, to participants at risk for weight loss (Simmons et al., 2003).

In general, weight loss is also associated with increased risk of mortality, functional ability, and transfer to a higher level of nursing home care (Murden & Ainslie, 1994; Ryan et al., 1995).

B. Summary of Differences Between the MDS 2.0 and the MDS 3.0 QM Definitions

The MDS 3.0 weight loss QM is based on the following specifications: Residents are counted in the numerator 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 have a weight loss of 5 percent or more of their body weight in the last month or a weight loss of 10 percent or more of their body weight over the last 6 months and were not on a physician-prescribed weight-loss regimen. Nursing facility residents with this condition have K0300=2 (weight loss) marked on the MDS 3.0 assessments, and the numerator counts the following types of assessments: quarterly, annual, significant change, or significant correction OBRA assessments (A0310A = 02, 03, 04, 05, 06); 14-, 30-, 60-, 90-day PPS assessments (A0310B = 02, 03, 04, 05); or discharge assessments (A0310F = 10, 11).

The underlying item definitions in the MDS 2.0 and 3.0 QM differ slightly. In the MDS 3.0, K0300 Weight Loss now has three response categories, with two new response categories referring to physician-prescribed weight loss. Only item response #2 for the K0300 item (below) is counted in the MDS 3.0 QM:

0. No or unknown.

1. Yes, on physician-prescribed weight-loss regimen.

2. Yes, not on physician-prescribed weight-loss regimen.

Residents are counted in the denominator 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 denominator consists of all assessments of long-stay residents, except admission (OBRA), PPS 5-day or PPS return/readmission assessments, and those for whom data on weight loss (K300) are missing. For public reporting, the score for any quarter is the simple average of the target quarter score and the quarter immediately prior to the target quarter. Facilities with fewer than 30 residents in the QM sample are excluded from public reporting because of low numbers and possible resident identification.

In the MDS 2.0 measure, residents were also excluded if they were receiving hospice care (P1A0 = checked) or hospice status was unknown (P1A0 = missing) on the target assessment or the most recent full assessment. This exclusion does not exist in the MDS 3.0 measure.

Reliability testing of MDS 3.0 data items used in calculating this measure as well as a comparison with the MDS 2.0 quality measures was conducted by RAND as part of the MDS 3.0 development process (Saliba & Buchanan, 2008). A representative sample of for-profit and not-for-profit facilities and hospital-based and free-standing facilities was recruited for the study, which included 71 community nursing facilities in eight States, 19 Veterans Affairs (VA) nursing facilities, and 1,390 nursing facility residents for the weight-loss quality measure. RAND compared the results on the nursing facility quality measures using the MDS 3.0 and the MDS 2.0, both at the individual-resident level and at the facility level (Saliba & Buchanan, 2008). At the resident level, the rate for weight loss using the MDS 2.0 was 8.3 percent and using the MDS 3.0 was 8.0 percent, with 96.1 percent agreement; the Kappa was 0.74, and the correlation was 0.74. Kappa is a statistical measure of inter-rater agreement ranging from 0.0 to 1.0. A rating of 0.74 is considered "substantial agreement." At the facility level, the MDS 2.0 rate was 8.6 percent and the MDS 3.0 rate was 8.3, with a correlation of 0.87.

Note also that the overall sample definitions also changed from the MDS 2.0 QM to the MDS 3.0 QM. In the MDS 2.0 QM, residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments. In the MDS 3.0 QM, the analogous sample is the long-stay residents, defined as a resident with more than 100 cumulative days in facility; the associated assessments may be any of the following (unless specifically excluded as described above): discharge; 14-, 30-, 60-, or 90-day PPS assessments; or quarterly, annual, significant change, or significant correction OBRA assessments.

C. Summary of Analyses

The following sections summarize the basic analyses performed on this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- number of assessments that are included and excluded from the numerator of the QM based on sample restrictions;
- whole sample prevalence of items that are related to the QM; and
- findings regarding variability, reportability, reliability, and validity.

4.15.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 1,171,863 (97.4 percent) assessments included in the denominator of QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay), and a total of 31,863 resident episodes (2.6 percent of the total number of long-stay resident episodes) were excluded from the denominator. ***Table 4.15-1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criteria. Note that one resident episode can meet multiple exclusion criteria. The number and percentage of resident episodes that met each exclusion criterion are as follows:

- A total of 9,741 (0.8 percent) assessments met exclusion criterion 1: missing data on items used to construct the numerator;
- and 22,402 (1.9 percent) assessments met exclusion criterion 2: assessments that were admission assessments, PPS 5-day, or readmission/return assessment.

Table 4.15-1
Target assessments used in the calculation of QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	1,171,863	97.4%
Excluded—Exclusion 1	9,741	0.8%
Excluded— Exclusion 2	22,402	1.9%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Missing Data

Exclusion 2 = Wrong target assessment type (admission assessment, PPS 5-day or readmission/return assessment)

a Column values may not add up to total because a resident episode can meet more than one exclusion criteria. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log)

Numerator Items – Distribution of Triggering Events

Table 4.15-2 describes the responses to the key MDS 3.0 item used to calculate this QM, K0300 Weight Loss, for all target assessments that were not excluded from the denominator due to missing data or wrong assessment. The overall sample prevalence of residents who lose too much weight while not on a physician-prescribed weight loss regiment is 7.1 percent (MDS 3.0 item K0300=2). Approximately 0.8 percent of target assessments are missing any data in this item. More than 91 percent report that there was no weight loss of this magnitude within the specified time frame (K0300 = 0). Less than 2 percent of target assessments indicated that this weight loss occurred while on a physician-prescribed weight-loss regimen (K0300 = 1).

Table 4.15-2
Responses to K0300, Loss of 5% or more in the last month or loss of 10% or more in last 6 months

Item	Entry	Frequency	Percentage
K0300	– = Missing Data	9,741	0.8%
K0300	0 = No or unknown	1,094,826	91.0%
K0300	1 = Yes, on a physician-prescribed weight loss regimen	13,428	1.1%
K0300	2 = Yes, not on physician-prescribed weight-loss regimen	85,731	7.1%
K0300	Total	1,203,726	—

Analysis date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_024_10.log)

4.15.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with perfect scores (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes) which may indicate a ceiling effect. ***Table 4.15-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) is 7.1 percent with a standard deviation of 4.6 percent. This QM has an interquartile range of 5.8 percent. About 5.7 percent of facilities have scores of 0 percent (perfect scores).

Table 4.15-3
QM score distribution for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)*

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
13,761	7.1%	4.6%	1.8%	3.8%	6.5%	9.6%	13.0%	5.7%	5.8%

*Measure averages quality measure over two quarters of data.

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure,

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_024_10.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30 long-stay residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examined the percentage of nursing homes that can report each measure (referred to as the QM reportability). ***Table 4.15-4*** shows the results of this analysis using MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 87.7 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) using MDS 3.0. This is an improvement over the MDS 2.0 measure, which 84.5 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may improve the reportability of long-stay QMs to some extent. Therefore, the reportability of short-stay QMs may be substantially improved after including the discharge assessment, but the impact on the long-stay QM should be less marked.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay), we report changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.15-4***. After excluding discharge assessments from the target assessment sample, reportability is largely unchanged for QM #0689 Percent of Residents Who Lose Too

Much Weight (Long Stay) (87.7 percent with discharge assessments versus 87.6 percent without discharge assessments).

Table 4.15-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator	1,171,863 (97.4%) [a]	1,167,097 (97.0%) [a]
Number (%) of resident episodes in the numerator	82,358 (6.8%) [a]	81,605 (6.8%) [a]
Number (%) of facilities able to report this QM	13,761 (87.7%) [b]	13,745 (87.6%) [b]

NOTES:

a Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay population (1,203,726).

b Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.15-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. Facility rankings were relatively unstable from one quarter to the next. In each of the three transitions in this analysis, fewer than half of facilities maintained the same rank (within one decile), and nearly four-fifths changed by less than three deciles. This is surprising: QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) is not a particularly low prevalence measure, and measure scores themselves were relatively stable (see next analysis). The variability in rank data for this measure merits monitoring in the future.

Table 4.15-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 4.6 percent). Overall, the average change in scores for this measure was a decline of 0.1 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (72.9 percent saw changes of less than one standard deviation), whereas approximately 1 percent saw changes of more than three standard deviations.

Table 4.15-5
Distribution of change in facility rank from one quarter to the next, QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	6,458 (47.9%)	6,537 (48.1%)	6,553 (48.2%)
Between 1 and 2 deciles	2,656 (19.7%)	2,684 (19.8%)	2,639 (19.4%)
Between 2 and 3 deciles	1,837 (13.6%)	1,754 (12.9%)	1,788 (13.2%)
More than 3 deciles	2,524 (18.7%)	2,610 (19.2%)	2,621 (19.3%)
Total	13,475	13,585	13,601

NOTES:

Total *n*'s reflect facilities that meet minimum requirements for public reporting this QM in both quarters.

Columns do not necessarily total 100.0% because of rounding.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 through Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.15-6
Change in facility scores from one quarter to the next, QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
13,601	-0.1%	4.2%	0.5%	2.2%	11.0%	34.8%	38.1%	10.7%	2.4%	0.5%

NOTES:

Number of facilities reporting (13,601) reflects facilities that meet minimum requirements for public reporting this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether State characteristics might be a source of facility score variation for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) and thus a potential threat to the measure's validity, we conducted a one-way analysis of variance (ANOVA). We also did a simple examination of the interquartile difference between the mean State-level scores across States. The proportion of variance in this measure explained by the State in which facilities are located is 3.3 percent and significant [$F(50, 13{,}710) = 9.3$, $p < .001$].[38] The difference between the mean State-level scores for States at the 25th percentile and the 75th percentile is 1.3 percentage points. Thus, although the majority of the variance in QM #0689 scores is due to factors other than geography, an exceptionally small and significant proportion of the variance is explained by the respective States in which nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns indicate that certain types of residents tend to have assessments with missing data in ways that influence the calculation of a quality measure, then that measure may be not be capturing processes and outcomes for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

In the fourth quarter of 2011, there were 32,143 residents who were excluded from the construction of this quality measure. After excluding the 22,402 residents (1.9 percent of the long-stay sample) with admission assessment or PPS 5-day or readmission/return assessments, the remaining 9,741 (0.8 percent) excluded residents had missing data on the weight loss item (K0300).

Table 4.15-7 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. Missing data are rare on items used in the calculation of this measure: the mean missing rate is 0.7 percent, and the median is 0 percent—values too small to represent a threat to the validity of the measure.

We further examined the possible relationship between missing data and QM scores. ***Table 4.15-8*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure. These distributions are essentially identical in their central tendency (near-zero for all) and in their variance. Further, there is no correlation between missing data and QM scores ($r = 0.0018$, *n.s*).

---

[38] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

Table 4.15-7
Distribution of facility-level missing rate for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,399	0.7%	2.9%	0%	0%	0%	0%	1.9%

NOTES:

n = number of facilities that have data for the numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request.log)

Table 4.15-8
Distribution of facility-level missing rate for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) by quartile of QM score

Quartile of QM score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	3,866	0.7%	2.7%	0%	0%	0%	0%	1.7%
26%–50%	3,803	0.8%	2.8%	0%	0%	0%	0%	2.1%
51%–75%	3,832	0.7%	2.5%	0%	0%	0%	0%	1.9%
76%–100%	3,831	0.6%	2.7%	0%	0%	0%	0%	1.6%

NOTES:

Total *n* = number of facilities that have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2) the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.15-9***. The mean nursing home–level QM score and distributions change negligibly before and after excluding discharge assessments from the set of eligible target assessments. The average facility score increases by less than one-tenth of 1 percent (with a median change of 0 percent) when discharge assessments are excluded. Thus, the use of the discharge assessment in the QM score calculations does not meaningfully alter the distribution of scores among facilities.

Table 4.15-9
Distribution of facility-level score change after excluding discharge assessment for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
13,745	0.0%	1.0%	-1.1%	0.0%	0.0%	0.1%	1.0%

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

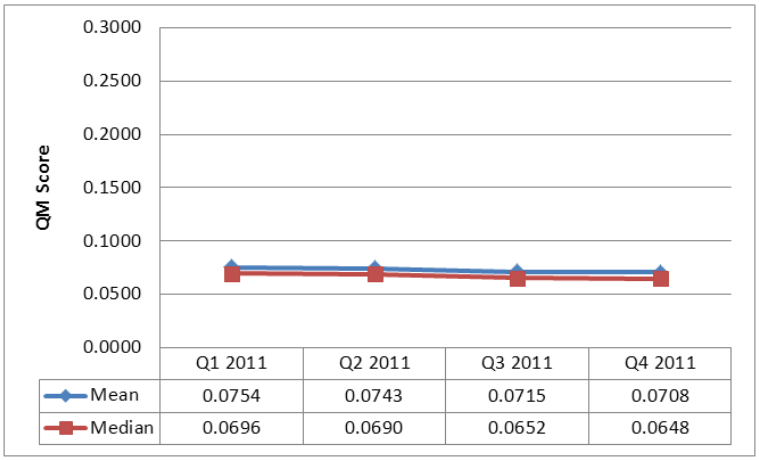
SOURCE: RTI analysis of MDS 3.0 data for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address this interest in whether seasonal variation might play a role in the score for QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay), we examined the national mean and median for this QM score between Quarter 1, 2011, and Quarter 4, 2011. The results are presented in ***Figure 4-15.1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, was 7.5 percent, 7.4 percent, 7.2 percent, and 7.1 percent. The quarterly national median was 7.0 percent, 6.9 percent, 6.5 percent, and 6.5 percent. Both mean and median decreased over all quarters. With only 12 months of data available, it is too early to ascertain whether the changes in national QM score over time are consistent with seasonal variation; this will need to be confirmed when multiple years of data become available.

Figure 4.15-1
Seasonal (quarterly) variation in QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay)



SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 1 - Quarter 4, 2011 \qm_quarter_1_2\complete\nh_024_10.log \qm_quarter_2_3\complete\nh_024_10.log \qm_quarter_3_4\complete\nh_024_10.log \qm_quarter_4_5\complete\nh_024_10.log

4.15.5 Risk Adjustment

QM #0689 Percent of Residents Who Lose Too Much Weight (Long Stay) was endorsed by NQF without denominator exclusion and model-based risk adjustment. Future analyses should examine the possibility of risk adjustment either through exclusions or indirect standardization with a particular focus on dementia and end-of-life status as potential exclusions or covariates.

4.16 QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

4.16.1 Summary of Findings

QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) reports the percentage of long-stay residents who have had symptoms of depression during the 2-week period preceding the MDS 3.0 target assessment date. This subsection presents descriptive analyses and test results for this QM. We found that 97.5 percent of long-stay resident episodes meet the denominator inclusion criteria and are included in the denominator. Based on these resident episodes, 87.8 percent of the facilities are able to report this QM (i.e., with 30 or more long-stay resident episodes included in the denominator). Among facilities able to report, the mean facility QM score is 7.3 percent. The QM score varies across facility with a standard deviation of 10.8 percent and interquartile range of 8.8 percent, suggesting acceptable variability of this QM to differentiate facilities with poor quality of care from those with good quality of care.

As reliability checks of this QM, we tracked the changes in QM score and in rank based on the QM score for each facility from quarter to quarter. The majority of facility score changes, either improvements or declines, were within one standard deviation with a very low proportion of changes being greater than three standard deviations. A similar pattern was found for facility rank changes. From quarter to quarter in 2011, more than 90 percent of facilities are with rank changes within three deciles.

We conducted several validity tests for this QM. First, we examined the frequency and distribution of missing data. Missing rate is 2.5 percent for the items used to construct this QM. The missing rate varies across facility, with some facilities having a missing rate greater than 7.5 percent. However, the missing rate is not significantly associated with the QM score, indicating that missing data should not pose a threat to validity of the QM. We also analyzed the potential geographic (State) variations in this QM. The State of the facility explains about 11.2 percent of the variation in this QM. The QM score is largely stable from quarter to quarter in 2011.

As reliability and validity tests support acceptable reliability and validity of this QM, QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) received full NQF endorsement on August 1, 2012.

4.16.2 Background and Introduction to QM

This outcome measure is based on data from MDS 3.0 assessments of nursing home residents, using those questions in MDS 3.0 that make up the Patient Health Questionnaire (PHQ-9) depression instrument. The PHQ-9 is based on the diagnostic criteria for a major depressive disorder in the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV). The resident interview measure will be used unless either there are three or more

missing subitems needed for calculation or the resident is rarely or never understood, in which cases the staff assessment measure will be calculated and used.

A. Summary of QM Background

Depression is a very expensive, complicating, and treatable factor for nursing facility residents. In a study of medically ill Medicare fee-for-service patients, patients with depression had significantly higher total health care costs than those without ($20,046 for patients with depression vs. $11,956 for patients without depression) (Unutzer et al., 2009). Higher costs were observed in participants with depression in every cost category except specialty mental health care, which accounted for less than 1 percent of total health care costs (Unutzer et al., 2009). The total economic cost of depression in the United States in calendar year (CY) 2000 was $83.1 billion, including $26.1 billion in direct medical costs (Greenberg et al., 2003). In the nursing facility environment, depression can be triggered by a number of elements of physical or cognitive decline, and by the circumstances of the nursing home residence itself (in addition to other causes), but can be underdiagnosed and undertreated (Simmons et al., 2003).

As summarized by (Saliba & Buchanan, 2008):

> Research conducted before the national implementation of the MDS demonstrated that the prevalence of major depression among cognitively intact or moderately impaired nursing facility residents was 20-25%. In addition, another 30% of residents had less severe, but nevertheless clinically significant depression (Parmelee, Katz, & Lawton, 1989). However, only about 10% of residents with recognized depression were treated. (Heston et al., 1992) More recent studies reveal that, despite an emphasis on depression in the MDS and associated quality indicators, as well as an almost 3 fold increase in the number of residents prescribed antidepressants, (Weintraub, Datto, Streim, & Katz, 2002) 34% of residents may have clinically significant depressive symptoms. (Datto et al., 2002)
>
> For the second quarter of 2008, the current measure ("Percent of Residents Who Have Become More Depressed or Anxious") based on MDS 2.0 data averages 14.9% nationally, with statewide averages ranging from 9.2% to 30%. Therefore, depression among the nursing home residents is a significant clinical issue.

This measure of depressed residents is directly related to an important component of overall health status. The illness does benefit from treatment, and untreated depression may contribute to a resident's decline. Better screening will increase the likelihood of treatment. Saliba and Buchanan report that 84 percent of the nurses in their study felt that the interview could inform facility care plans, and that 86 percent reported that even in the limited number of residents assessed, the interview items provided new insights into resident mood (Saliba & Buchanan, 2008). They also reported that for the PHQ-9 staff observation version, 90 percent felt that staff detection and communication about mood disorder might improve if they learned to watch for these signs and symptoms. This is an important finding, given that geriatric depression is a common but frequently unrecognized or inadequately treated condition in the elderly population.

B. Summary of Differences between MDS 2.0 and MDS 3.0 QM Definitions

Residents are counted in the numerator 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. Using the PHQ-9 items in the MDS 3.0, for the Resident Interview Measure (Item D0200), the numerator is based on the residents' little interest in doing things or feelings of depression, and the total sum severity score (D0300) on the most recent MDS assessment in the selected quarter (which may be discharge assessments; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments).

A resident with little interest or pleasure in doing things half or more days over the last 2 weeks (D0200A2 = 02, 03); or feeling down, depressed, or hopeless half or more days over the last 2 weeks (D0200B2 = 02, 03); can be eligible for inclusion in the numerator in one of two ways for the MDS 3.0: the Resident Mood Interview or Staff Assessment of Resident Mood.

The total severity score (D0300) reflects resident responses to questions asking about the frequency of nine symptoms over the last 2 weeks, including interest, mood, energy, appetite, self-value, ability to concentrate, change in responsiveness, or patience. The Staff Assessment Measure (Item D0500) is similar, except the judgment is being made by observers rather than the residents themselves. A total score is calculated from Column 2, Symptom Frequency, for items in D0200.

When the Resident Mood Interview is conducted, the resident must have a score of two or greater for either D0200A or D0200B *and* a score of two or more for five of the following items D0200A to D0200I (i.e., the sum for D0300 is equal or greater than 10 and less than or equal to 27). When the Staff Assessment for Resident Mood is necessary, the resident must have a score of two or greater for either D0500A or D0500B *and* a score of two or more for five of the following items D0500A to D0500I (i.e.: the sum for D0600 is equal or greater than 10 and less than or equal to 30). The Staff Assessment of mood (items D0500) is used if a long-stay resident is missing data for three or more of the subitems of data elements D0200 for the Resident Assessment *and* has valid data for seven or more of subitems A through I of item D0500 for the Staff Assessment.

Residents are counted in the denominator 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. The target population for the denominator is the total number of all long-stay residents in the nursing facility who have received an MDS assessment (which may be discharge; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments) during the selected quarter (3-month period) and who do not meet the exclusion criteria.

The following items exclude residents from the denominator:

- B0100 = 1 or missing (Comatose), OR
- The resident is not included in the numerator (the resident did not meet the depression symptom conditions in the numerator) AND both of the following are true:
  - D0200A2 is missing, or D0200B2 is missing, or D0300 is missing or Symptom Frequency is blank for 3 or more items (indicating that interview could not be completed).
  - D0500A2 is missing, D0500B2 is missing, or D0600 is missing.

The transition from MDS 2.0 to MDS 3.0 yielded differences in how the measure was defined and how data to calculate the measure were collected. First, in regard to measure definition, the MDS 2.0 measure described worsening of mood, as measured by a Mood Scale Score assessed by staff, between the target assessment and prior assessment. The implication is that the MDS 2.0 measure reflected only those residents whose depression or anxiety worsened. In addition, the MDS 2.0 data combined two separate conditions (depression and anxiety), as well as situations that may result from other causes entirely: distress, crying/tearfulness, motor agitation, leaves food uneaten, repetitive health complaints, repetitive/recurrent verbalizations, negative statements, and mood symptoms not easily altered. In contrast, the MDS 3.0 measure uses an assessment in one time period only, and is an indication of the presence of symptoms of major depression.

Second, in regard to data elements, the MDS 3.0 replaces the Mood Scale Score with a resident interview version and a staff assessment version of the PHQ-9, which is based on the diagnostic criteria for a major depressive disorder in the DSM-IV. The resident interview version, which is preferred, is a major change from the MDS 2.0. The numerator is based on the two specific items of the PHQ-9 and the total sum severity score (D0300) on the most recent MDS assessment in the selected quarter. The total severity score reflects resident responses to questions asking about frequency of nine symptoms over the last 2 weeks, including interest, mood, energy, appetite, self-value, ability to concentrate, change in responsiveness or patience. If a resident is unable to complete the interview items, then staff complete the Staff Assessment Measure (D0500) items and the numerator is based on the total sum score (D0600).

Third, the overall sample definitions also changed from MDS 2.0 to MDS 3.0. For the MDS 2.0, residents were included in the chronic care measures if they had a full or quarterly MDS in the target quarters, and measures were calculated based on just these two types of assessments. For the MDS 3.0, the analogous sample is for long-stay residents, defined as residents with more than 100 cumulative days in a facility, and assessments may be discharge; 5-, 14-, 30-, 60-, 90-day, or readmission/return PPS assessments; or admission, quarterly, annual, significant change, or significant correction OBRA assessments.

There were several reasons for changing the way this quality measure was defined and the data elements used in its calculation. First, the list of 15 observed indicators of depression has poor sensitivity for identifying individuals with depressive symptoms or depression (Anderson, Buckwalter, Buchanan, Maas, & Imhof, 2003; Horgas & Tsai, 1998; McCurren, 2002; Snowden, Sato, & Roy-Byrne, 2003; Teresi, Abrams, Holmes, Ramirez, & Eimicke, 2001). Second, only 22 percent of nurses in their survey reported that the MDS 2.0 mood items are easy to complete accurately (Saliba & Buchanan, 2008).

Saliba and Buchanan (2008) analyzed the reliability and validity of the PHQ-9 in a sample of 71 community and 19 VA nursing facilities distributed throughout the United States. Residents were selected in these facilities to capture a representative sample of short- and long-stay residents, and in order to maximize the number of MDS 2.0 items assessed, the selection algorithms included a strong preference for capturing cases scheduled for MDS 2.0 admission assessment. They compared the PHQ-9 with two “gold-standard” measures: the Modified Schedule for Affective Disorders and Schizophrenia (m-SADS), and the Cornell Depression Scale, using kappas as measures of concordance. These alternative measures were also administered to a subsample of residents, and the results compared with those from the PHQ-9.

Saliba and Buchanan (2008) found that reliability was excellent: the average kappa between gold-standard nurses for the PHQ-9 resident interview was 0.935, and between gold-standard and facility nurses it was 0.968. The proposed quality measure is a ratio constructed from these measures and is therefore reliable. The kappa for the PHQ-9 and the m-SADS was very good. The correlation between the PHQ-9 and the Cornell depression scale was 0.63.

C. Summary of Analyses

The following sections summarize the basic analyses performed on this QM using MDS 3.0 data for the QM for Quarter 4, 2011(unless otherwise specified). Three general areas are addressed:

- number of assessments that are included and excluded from the numerator of the QM based on sample restrictions;
- whole sample prevalence of items that are related to the QM; and
- findings regarding variability, reportability, reliability, and validity.

4.16.3 Descriptive Statistics

QM Numerator/Denominator Selection/Exclusions

Overall, in Quarter 4, 2011, there were 1,173,873 (97.5 percent) assessments included in the denominator of QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay), and a total of 29,853 resident episodes (2.5 percent of the total number of long-stay resident episodes) were excluded from the denominator. ***Table 4.16-1*** shows the proportion of resident episodes excluded from the denominator for each of the measure's exclusion criteria. Note that one resident episode can meet multiple exclusion criteria. The number and percentage of resident episodes that met each exclusion criterion are as follows:

- A total of 29,853 (2.5 percent) assessments met exclusion criterion 1: missing data on items used to construct the numerator; and
- 4,554 (0.3 percent) assessments met exclusion criterion 2: comatose status.

Table 4.16-1
Target assessments used in the calculation of QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

Resident episodes in the reporting period	Frequency[a]	Percentage[a]
Included	1,173,873	97.5%
Excluded—Exclusion 1	29,853	2.5%
Excluded— Exclusion 2	4,554	0.3%
Total number of long-stay resident episodes	1,203,726	—

NOTES:

Exclusion 1 = Missing Data

Exclusion 2 = Comatose

[a] Column values may not add up to total because a resident episode can meet more than one exclusion criterion. Percentage column reflects percentage of target assessments in each category out of total resident episodes in long-stay population.

Analysis date: 5/25/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db127_request_v1.log).

Numerator Items – Distribution of Triggering Events

Of all resident target assessments that could be included in the numerator for this QM, ***Table 4.16-2*** shows the prevalence of responses for each item used to calculate this QM across all assessments. For D0200A2 or D0500A2, the frequency of having little interest or pleasure in doing things, and D0200B2 or D0500B2, the frequency of feeling down, depressed, or hopeless, the categories correspond to the following:

In the past 2 weeks, resident has been bothered:

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

A [-] or [^] indicates, respectively, that the item was missing or skipped.

In just over one-quarter of target assessments, the Resident Mood Interview was not completed, and instead the Staff Assessment of Resident Mood was used. It also shows that of residents completing the Resident Mood Interview, only 4.31 percent had a score between 10 and 27 that would partially qualify them for the numerator. Similarly, of residents who were assessed by staff, only 3.8 percent had a score that would qualify them for the numerator.

4.16.4 Findings from Testing

Variability

A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. Of particular interest is whether there is evidence of nursing homes clustering along the distribution of QM scores, which can be indicated by the width of the interquartile range (the difference between the 75th and 25th percentiles), or shown by the percentage of nursing homes with perfect scores (i.e., 0 percent triggering rates for measures capturing negative outcomes and 100 percent for measures reflecting positive outcomes), which may indicate a ceiling effect. ***Table 4.16-3*** shows the results of this analysis using the QM score for Quarter 4, 2011. The mean score for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) is 7.3 percent with a standard deviation of 10.8 percent. This QM has an interquartile range of 8.8 percent. About 24.3 percent of facilities have scores of 0 percent (perfect scores).

Table 4.16-2
Frequency of items that count toward numerator of QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

Item	Entry	Frequency	Percentage	Cumulative %
D0200A2	— = missing	46,939	3.9%	3.9%
D0200A2	0 = Never or 1 day	717,978	59.7%	63.6%
D0200A2	1 = 2-6 days (several days)	47,606	4.0%	67.5%
D0200A2	2 = 7-11 days (half or more of the days)	27,754	2.3%	69.8%
D0200A2	3 = 12-15 days (nearly every day)	34,525	2.9%	72.7%
D0200A2	^ = skipped	328,924	27.3%	100.0%
D0200A2	Total	1,203,726	—	—
D0200B2	— = missing	47,499	4.0%	4.0%
D0200B2	0 = Never or 1 day	601,802	50.0%	53.9%
D0200B2	1 = 2-6 days (several days)	124,750	10.4%	64.3%
D0200B2	2 = 7-11 days (half or more of the days)	54,931	4.6%	68.9%
D0200B2	3 = 12-15 days (nearly every day)	47,848	4.0%	72.8%
D0200B2	^ = skipped	326,896	27.2%	100.0%
D0200B2	Total	1,203,726	—	—
D0300	— (missing)	45,772	3.80%	3.80%
D0300	0	369,741	30.71%	34.51%
D0300	1	88,641	7.36%	41.87%
D0300	2	79,302	6.59%	48.46%
D0300	3	74,741	6.21%	54.67%
D0300	4	46,071	3.83%	58.50%
D0300	5	32,418	2.70%	61.20%
D0300	6	33,925	2.82%	64.02%
D0300	7	19,047	1.58%	65.60%
D0300	8	15,183	1.27%	66.87%
D0300	9	15,023	1.25%	68.12%
D0300	10	12,537	1.04%	69.16%
D0300	11	8,036	0.67%	69.83%
D0300	12	8,791	0.73%	70.56%

(continued)

Table 4.16-2 (continued)
Frequency of items that count toward numerator of QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

Item	Entry	Frequency	Percentage	Cumulative %
D0300	13	4,310	0.36%	70.92%
D0300	14	3,563	0.30%	71.22%
D0300	15	4,139	0.34%	71.56%
D0300	16	1,887	0.16%	71.72%
D0300	17	1,501	0.12%	71.84%
D0300	18	1,926	0.16%	72.00%
D0300	19	743	0.06%	72.06%
D0300	20	534	0.04%	72.10%
D0300	21	852	0.07%	72.17%
D0300	22	207	0.02%	72.19%
D0300	23	161	0.01%	72.20%
D0300	24	342	0.03%	72.23%
D0300	25	44	0.00%	72.23%
D0300	26	26	0.00%	72.23%
D0300	27	105	0.01%	72.24%
D0300	99	64,940	5.39%	77.63%
D0300	^ = skipped	269,218	22.37%	100.00%
D0300	Total	1,203,726	—	—
D0500A2	— = missing	20,422	1.7%	1.7%
D0500A2	0 = Never or 1 day	271,459	22.6%	24.3%
D0500A2	1 = 2-6 days (several days)	19,904	1.7%	25.9%
D0500A2	2 = 7-11 days (half or more of the days)	18,709	1.6%	27.5%
D0500A2	3 = 12-15 days (nearly every day)	44,882	3.7%	31.2%
D0500A2	^ = skipped	828,350	68.8%	100.0%
D0500A2	Total	1,203,726	—	—
D0500B2	— = missing	20,054	1.7%	—
D0500B2	0 = Never or 1 day	285,006	23.7%	—

(continued)

Table 4.16-2 (continued)
Frequency of items that count toward numerator of QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

Item	Entry	Frequency	Percentage	Cumulative %
D0500B2	1 = 2-6 days (several days)	28,822	2.4%	—
D0500B2	2 = 7-11 days (half or more of the days)	19,349	1.6%	—
D0500B2	3 = 12-15 days (nearly every day)	22,145	1.8%	—
D0500B2	^ = skipped	828,350	68.8%	—
D0500B2	Total	1,203,726	—	—
D0600	— (missing)	21,349	1.8%	1.8%
D0600	0	143,817	11.9%	13.7%
D0600	1	21,290	1.8%	15.5%
D0600	2	20,943	1.7%	17.2%
D0600	3	34,122	2.8%	20.1%
D0600	4	17,345	1.4%	21.5%
D0600	5	14,205	1.2%	22.7%
D0600	6	21,979	1.8%	24.5%
D0600	7	10,422	0.9%	25.4%
D0600	8	9,664	0.8%	26.2%
D0600	9	13,039	1.1%	27.3%
D0600	10	8,297	0.7%	28.0%
D0600	11	6,300	0.5%	28.5%
D0600	12	9,461	0.8%	29.3%
D0600	13	4,118	0.3%	29.6%
D0600	14	3,595	0.3%	29.9%
D0600	15	5,341	0.4%	30.3%
D0600	16	2,023	0.2%	30.5%
D0600	17	1,716	0.1%	30.7%
D0600	18	2,697	0.2%	30.9%
D0600	19	879	0.1%	30.9%
D0600	20	675	0.1%	31.0%
D0600	21	1,127	0.1%	31.1%

(continued)

Table 4.16-2 (continued)
Frequency of items that count toward numerator of QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

Item	Entry	Frequency	Percentage	Cumulative %
D0600	22	258	0.0%	31.1%
D0600	23	239	0.0%	31.1%
D0600	24	296	0.0%	31.2%
D0600	25	48	0.0%	31.2%
D0600	26	42	0.0%	31.2%
D0600	27	63	0.0%	31.2%
D0600	28	12	0.0%	31.2%
D0600	29	3	0.0%	31.2%
D0600	30	11	0.0%	31.2%
D0600	^ = skipped	828,350	68.8%	100.0%
D0600	Total	1,203,726	—	—

NOTE: D0200A2 = Little interest of pleasure in doing things: Symptom frequency (Resident Interview); D0200B2 = Feeling down, depressed, or hopeless: Symptom Frequency (Resident Interview); D0300 = PHQ-9© Total Severity Score (Resident Interview); D0500A2 = Little interest of pleasure in doing things: Symptom frequency (Staff Assessment); D0500B2 = Feeling down, depressed, or hopeless: Symptom Frequency (Staff Assessment); D0600 = PHQ-9© Total Severity Score (Resident Interview).

Analysis date: 4/17/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_025_10.log)

Table 4.16-3
QM score distribution for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

n	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with "perfect scores"	Interquartile range
13,775	7.3%	10.8%	0%	0.7%	3.8%	9.5%	18.2%	24.3%	8.8%

Analysis date: 4/17/2012

NOTES:

n = number of facilities that meet minimum requirements for public reporting this quality measure.

QM scores are reported at the facility level.

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\qm_quarter_4_5\complete\nh_025_10.log)

Reportability

For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 30

long-stay residents who qualified for the denominator of this measure) after applying measure exclusion criteria. We therefore examined the percentage of nursing homes that can report each measure (referred to as the QM reportability). ***Table 4.16-4*** shows the results of this analysis using of MDS 3.0 QMs calculated for Quarter 4, 2011. After applying measure exclusion criteria, 87.8 percent of facilities had sufficient sample size to meet minimum requirements for public reporting QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) using MDS 3.0. This is an improvement over the MDS 2.0 measure, which 85.3 percent of facilities were able to report in Quarter 1, 2006 (Brega, Goodrich, Nuccio, et al., 2008).

The inclusion of the discharge assessment is an important change from the MDS 2.0 to the MDS 3.0 measure and may increase the number of resident episodes eligible to be included in the QM samples. The inclusion of the discharge assessment should capture short-stay residents who are discharged prior to their 14-day assessment who would have been missed under the MDS 2.0 sample specification. The inclusion of the discharge assessment also allows for more complete data on residents who are discharged between the 14- and 30-day assessments. As described in ***Table 2-2*** in Section 2, just over 75 percent of target assessments for the short-stay sample for Quarter 4, 2011, were discharge assessments, and that 8.8 percent of target assessments for the long-stay sample in the same quarter were discharge assessments. This suggests that the inclusion of the discharge assessment may improve the reportability of long-stay QMs to some extent. Therefore, the reportability of short-stay QMs may be substantially improved after including the discharge assessment, but the impact on the long-stay QM should be less marked.

To evaluate the impact on reportability of including the discharge assessment in the set of target assessments eligible for constructing QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay), we report changes in the number of facilities able to report this QM when discharge assessments are not included in the target assessment sample. The results are presented in ***Table 4.16-4.*** After excluding discharge assessments from the target assessment sample, reportability is largely unchanged for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) (87.82 percent with discharge assessments vs. 87.78 percent without discharge assessments).

Table 4.16-4
Change in reportability based on inclusion or exclusion of discharge assessments for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

Key value	Discharge assessment included	Discharge assessment excluded
Number (%) of resident episodes in the denominator	1,173,873 (97.5%) [a]	1,186,913 (98.6%) [a]
Number (%) of resident episodes in the numerator	84,341 (7.0%) [a]	85,541 (7.1%) [a]
Number (%) of facilities able to report this QM	13,775 (87.8%) [b]	13,768 (87.8%) [b]

NOTES:

[a] Percentage of resident episodes included in the denominator or numerator for calculating this QM is out of the total number of resident assessments in the long-stay population (1,203,726).

[b] Percentage of facilities is out of the total number of facilities with at least one short-stay or long-stay resident (15,686).

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

Reliability

Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. We examined the extent to which relative facility rank changed on this quality measure from quarter to quarter in 2011. We also examined the change in facility scores themselves from Quarter 3 to Quarter 4 of 2011.

Table 4.16-5 shows how facility ranks for this quality measure changed from quarter to quarter in 2011. For each pair of quarters, large changes in ranking were rare: approximately two-thirds of facilities stayed in approximately the same rank, and relatively few (no more than 8.3 percent) facilities saw their rank change by more than three deciles.

Table 4.16-5
Distribution of change in facility rank from one quarter to the next, QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

Magnitude of shift	Quarter 1 to Quarter 2, 2011: *n* (%)	Quarter 2 to Quarter 3, 2011: *n* (%)	Quarter 3 to Quarter 4, 2011: *n* (%)
Within 1 decile	8,657 (64.5%)	8,956 (66.0%)	9,196 (67.6%)
Between 1 and 2 deciles	2,470 (18.4%)	2,189 (19.1%)	2,074 (15.3%)
Between 2 and 3 deciles	1,283 (9.6%)	1,295 (9.6%)	1,250 (9.2%)
More than 3 deciles	1,012 (7.6%)	1,121 (8.3%)	1,082 (8.0%)
Total	13,423	13,561	13,602

NOTES:

Total *n*'s reflect facilities that meet minimum requirements for public reporting this QM in both quarters.

Analysis date: 6/15/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 through Quarter 4, 2011 (\quarter_4_5\db155_request\db155_request.log)

Table 4.16-6 presents the results from an analysis of the change in facility scores for this quality measure between Quarters 3 and 4 of 2011. Because quality measures vary in their central tendency and in their variance, it is difficult to compare absolute changes in scores across measures. For this analysis, we normalized the scores and calculated proportions of facilities whose scores remained approximately the same and those that changed by more than one, two, or three standard deviations (standard deviation = 10.8 percent). Overall, the average change in scores for this measure was an increase of 0.4 percent. The vast majority of facilities saw no meaningful change in their score from quarter to quarter (84 percent saw changes of less than one standard deviation), whereas fewer than 2 percent of facilities saw changes of more than three standard deviations.

Table 4.16-6
Change in facility scores from one quarter to the next, QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

n of facilities reporting	Mean QM score change	SD of QM score change	Proportion of facilities that declined by more than three standard deviations	Proportion of facilities that declined by between two and three standard deviations	Proportion of facilities that declined by between one and two standard deviations	Proportion of facilities that declined by less than one standard deviation	Proportion of facilities that improved by less than one standard deviation	Proportion of facilities that improved by between one and two standard deviations	Proportion of facilities that improved by between two and three standard deviations	Proportion of facilities that improved by more than three standard deviations
13,602	0.4%	5.3%	0.4%	1.3%	6.1%	55.8%	28.2%	5.4%	1.4%	1.4%

NOTES:

Number of facilities reporting (13,602) reflects facilities that meet minimum requirements for public reporting this QM in both Quarter 3 and Quarter 4, 2011.

Analysis date: 6/18/2012

SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 3 and Quarter 4, 2011 (\quarter_4_5\db154_request\db154_request_combadj_reportable_v4_new.log)

Validity

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a threat to the validity of the measure. To explore the question of whether State characteristics might be a source of facility score variation for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) and thus a potential threat to the measure's validity, we conducted a one-way analysis of variance (ANOVA). We also did a simple examination of the interquartile difference between the mean State-level scores across States. The proportion of variance in this measure explained by the State in which facilities are located is 11.2 percent and significant [$F(51, 13{,}723) = 34.0$, $p < .001$]. [39] The difference between the mean State-level scores for States at the 25th percentile and the 75th percentile is 3.9 percentage points. Thus, although the majority of the variance in QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) scores is due to factors other than geography, a small (but large relative to other measures for which this analysis was conducted) and significant proportion of the variance is explained by the respective States in which nursing facilities are located. This feature of the measure warrants future monitoring.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns indicate that certain types of residents tend to have assessments with missing data in ways that influence the calculation of a quality measure, then that measure may not be capturing processes and outcomes for the intended population, thus inflating or suppressing QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised.

In the fourth quarter of 2011, there were 29,853 (2.5 percent of the long-stay population) residents excluded from the construction of this measure. All of the excluded residents had missing data on MDS items related to mood. In addition, 4,554 of those residents were comatose at the time of their target assessment (a possible explanation for missing data on mood items, though contrary to the instructions given in the MDS 3.0 specifications).

Table 4.16-7 provides summary statistics for an analysis of the distribution of missing data rates for facilities reporting on this measure. The mean facility-level missing-data rate for items used to calculate this measure is 2.6 percent. However, at least half of facilities have no missing data on those items, and at the rate at the 90th percentile of facilities is 7.5 percent, indicating that relatively few facilities may be skewing the distribution. It may be valuable to assess relevant characteristics of those outlying facilities in order to test for systematic bias.

[39] SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db160_request)

Table 4.16-7
Distribution of facility-level missing rate for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

n	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
15,399	2.6%	5.5%	0%	0%	0%	3.1%	7.5%

NOTES:

n = number of facilities that have data for the numerator and denominator of this QM before assessment exclusion criteria are applied; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/5/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db177_request\db177_request.log)

We further examined the possible relationship between missing data and QM scores. ***Table 4.16-8*** shows the distributions of facility-level missing rates stratified by quartiles of QM scores for this measure. These distributions are nearly identical in central tendency (between 2 percent and 3 percent) and in variance. Further, there is no correlation between missing data and QM scores (r = -0.0144, *n.s.*).

Table 4.16-8
Distribution of facility-level missing rate for QM #0690 Percent of Residents Who Have Depressive Symptoms by quartile of QM score (Long Stay)

Quartile of QM Score	*n*	Mean	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
0%–25%	4,139	2.8%	6.6%	0%	0%	0%	2.9%	7.7%
26%–50%	3,557	2.3%	3.9%	0%	0%	5.9%	2.8%	7.0%
51%–75%	3,863	2.6%	4.5%	0%	0%	8.1%	3.3%	7.4%
76%–100%	3,832	2.5%	4.7%	0%	0%	0%	3.4%	7.5%

NOTES:

Total *n* = number of facilities that have data for the numerator and denominator of this QM; facilities are included regardless of whether they meet the minimum sample size for reporting.

Analysis date: 7/26/2012

SOURCE: RTI analysis of MDS 3.0 episode file for Quarter 4, 2011 (\quarter_4_5\db181_request\db181_request.log)

Change in Scores after Excluding Discharge Assessment. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2) the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. The results are presented in ***Table 4.16-9.*** Change in the mean nursing home–level QM score and distribution after excluding discharge assessments from the set of eligible target assessments is negligible. The average facility score decreases by less than one-tenth of 1 percent (with a median change of 0%) when discharge assessments are excluded. Thus, the use of the discharge assessment in the QM score calculations does not meaningfully alter the distribution of scores among facilities.

Table 4.16-9
Distribution of facility-level score change after excluding discharge assessment for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)

n	Mean score change	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
13,768	0.0%	0.9%	-0.7	0.0	0.0	0.0	0.9

NOTES:

The number of facilities in this analysis represents all facilities that could report this QM (i.e., meet minimum sample size for reporting) before and after excluding discharge assessments.

Analysis date: 5/29/2012

SOURCE: RTI analysis of MDS 3.0 data for Quarter 4, 2011 (\quarter_4_5\db135_request\db135_request.log)

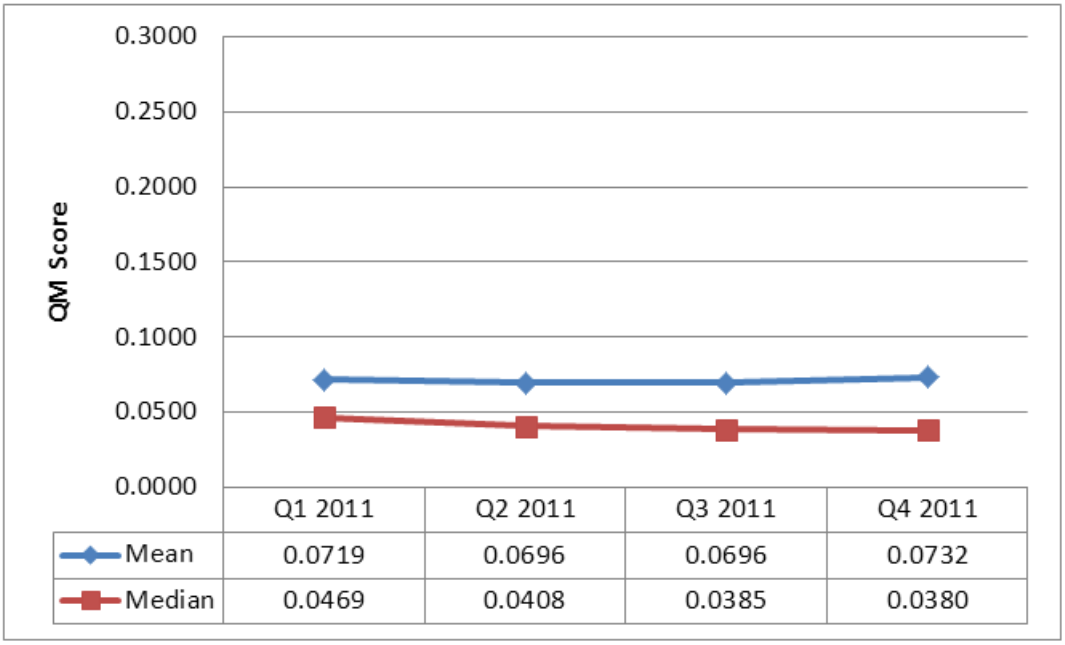
Seasonal Variation

Another potential threat to the validity of a QM is seasonal variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, this suggests a threat to the validity of the measure because it is being influenced by factors outside of the nursing home's control. To address this interest in whether seasonal variation might play a role in the score for QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay), we examined the national mean and median for this QM score between Quarter 1, 2011 and Quarter 4, 2011. The results are presented in ***Figure 4.16-1***.

The national-level mean for this measure for each quarter between Quarter 1, 2011, and Quarter 4, 2011, was 7.2 percent, 7.0 percent, 7.0 percent, and 7.3 percent. The quarterly national median was 4.7 percent, 4.1 percent, 3.9 percent, and 3.8 percent. Although the mean decreased from Quarter 1, 2011, to Quarter 2, 2011, remained constant from Quarter 2, 2011, to Quarter 3, 2011, and increased from Quarter 3, 2011, to Quarter 4, 2011, the median decreased over all four quarters. With only 12 months of data available, it is too early to ascertain whether

the changes in national QM score over time are consistent with seasonal variation; this will need to be confirmed when multiple years of data become available.

Figure 4.16-1
Seasonal (quarterly) variation in QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay)



SOURCE: RTI analysis of MDS 3.0 episode files for Quarter 1 - Quarter 4, 2011 \qm_quarter_1_2\complete\nh_025_10.log \qm_quarter_2_3\complete\nh_025_10.log qm_quarter_3_4\complete\nh_025_10.log \qm_quarter_4_5\complete\nh_025_10.log

4.16.5 Risk Adjustment

QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) is risk adjusted through denominator exclusions. Residents who are comatose are excluded from the denominator. Future analyses should examine the possibility of risk adjustment with a particular focus on cognitive status as potential risk adjusters.

SECTION 5
CONCLUSIONS

This report details analyses of MDS 3.0 nursing home quality measures in terms of reportability, variability, reliability, validity, and risk adjustment. The analyses in this report are based primarily on the episode file that RTI created for calculating the QMs for the fourth quarter of 2011. The episode file primarily comprises MDS 3.0 assessments from July 1 to December 31 (Quarters 3 and 4, 2011) but also includes some assessments from the second quarter of 2011 for some long-stay residents. Coding of the quality measures was based on the specifications detailed in the MDS 3.0 Quality Measures User's Manual version 5.0. Analyses regarding short-stay and long-stay definitions are also presented.

Short-Stay/Long-Stay Definitions. The parameters defining the short- and long-stay residents changed from the MDS 2.0 to the MDS 3.0. In the use of the MDS 2.0, residents were included in the post-acute care (PAC) measures if they had a 14-day PPS MDS assessment in the target quarters, and measures were calculated based only on 14-day PPS assessments. In the MDS 3.0, short-stay and long-stay were redefined: residents with 100 or fewer cumulative days in facility (CDIF) would be categorized as short-stay residents and those with more than 100 CDIF would be categorized as long-stay residents. Analyses indicate that most facilities provide services to a mix of short-stay and long-stay residents. One consequence of the new MDS 3.0 definition of short stay is that some residents who are classified as short-stay may become long stay in the next quarter as they stay in the nursing homes and eventually accumulate more than 100 CDIF. For example, 10.6 percent of residents classified as short-stay in Quarter 3 of 2011 were reclassified as long-stay in Quarter 4 of 2011. We investigated the impact of including these "early long-stay" residents in constructing short-stay QMs. These short-stay QM scores were recalculated after excluding early long-stay and then compared with the scores calculated on the overall short-stay population (i.e., including early long-stay). The results from all four short-stay QMs indicate little impact of including early long stay on the QM scores.

Findings from Testing. The 16 nursing home QMs were submitted for NQF endorsement. This process required an analytic design for investigating quality measures that comprises five general areas: variability, reportability, reliability, validity, and risk-adjustment analyses. For validity testing, RTI conducted several analyses, including correlation analyses between paired/grouped measures, the impact of the State in which the facility is located on the QM scores, the impact of missing data on QM scores, and seasonal (quarterly) variation.

<u>Variability.</u> A well-designed QM should capture the actual variability in quality of care that exists across nursing homes and should discriminate among nursing homes that are performing well and those that are performing poorly. In order to test variability, we examined the spread of the distributions of scores by calculating their interquartile ranges and looked for ceiling effects by calculating the proportions of facilities with perfect scores. The interquartile ranges for the QMs range from 2.9 percent for QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) to 24.8 percent for QM #0685 Percent of Low Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay). The percentage of facilities with "perfect scores" ranges from 0.3 percent for QM #0685 Percent of Low Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) to 50.5 percent for QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay). QMs with narrow

interquartile ranges tend to be associated with high percentage of facilities with perfect scores and vice versa.

Reportability. For a QM to be useful for public reporting, the majority of nursing homes should have sufficient sample size to meet minimum requirements for public reporting (i.e., had at least 20 short-stay residents who qualified for the denominator of the short-stay QMs; at least 30 long-stay residents who qualified for the denominator of the long-stay QMs) after applying measure exclusion criteria. The majority of the MDS 3.0 QMs had relatively high percentages of facilities who met the minimum sample size reporting thresholds. Given minimum sample size requirements, the percentage of facilities able to report a certain QM ranges from 42.6 percent for QM #0685 Percent of Low Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) to 88.4 percent for three QMs. These three QMs are QM #0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay), QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccination (Long Stay), and QM #0687 Percent of Residents Who Were Physically Restrained (Long Stay). Aside from QM #0685 Percent of Low Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) and QM #0678 Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay) (which had a reportability of 68.6 percent), the remaining measures each had a rate of reportability greater than 70 percent.

Reliability. Reliability refers to the consistency of results obtained when a measurement is repeated or the vulnerability of a measurement to random error. Dramatic changes in facility QM scores from one quarter to the next are more likely an indicator of measure instability than of great improvement or decline in facility performance. The MDS 3.0 QM scores between the two quarters analyzed resulted in fairly consistent facility ratings. Most facilities remained the same QM scores or slightly improved from Quarter 3 to Quarter 4 of 2011. For each QM, at least 71.2 percent (QM #0685 Percent of Low Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) and as many as 88.6 percent (QM #0683 Percent of Residents Assessed and Appropriately Given the Pneumococcal Vaccine [Long Stay]) of facilities had the same score (within one standard deviation) in Quarter 3 and Quarter 4 of 2011.

Validity. Four principal strategies were used to examine the validity of the 16 quality measures:

Correlations. Although nursing home quality measures have historically illustrated low correlations among measures, a common strategy used to evaluate validity is to examine the facility percentile ranking correlation among groups of measures that capture related clinical care processes (convergent validity; i.e., their percentile ranking on any of these measures, should be correlated). All of the measure pairs were significantly correlated except for the pairing of QM #0685 Percent of Low-Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) and QM #0686 Percent of Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay).

Variation by State. For a quality measure to be valid, variation observed in the distribution of the measure should be attributable to differences in the underlying quality of the facilities being evaluated. If a measure is subject to variation caused by other factors outside of the control of facilities, such as State-level payment policies or demographics, this can be a

threat to the validity of the measure. There was a significant effect of State for each of the 16 quality measures. However, State location accounted for less than 5 percent of the variance in quality measure scores for 8 measures and greater than 10 percent for just 2: QM #0685 Percent of Low Risk Residents Who Lose Control of Their Bowel or Bladder (Long Stay) at 16.2 percent, and QM #0690 Percent of Residents Who Have Depressive Symptoms (Long Stay) at 11.2 percent. It is unclear from these analyses whether these relatively large effect sizes stem from the geographic differences between States (for example, northern States might have more residents suffering from seasonal affective disorder, thus driving up scores on the depression measure) or if State-level policies have a disproportionate impact on care processes relating to these two measures.

Missing Data. Missing data represent a potential threat to the validity of a quality measure. If patterns indicate that similar types of residents and/or assessments have missing data, the measure may not be capturing processes and outcomes for the intended population, thus inflating or suppressing facility-level QM scores. Further, if missing data rates are systematically dissimilar across facilities, then the ability to compare facilities on certain measures may be compromised. The impact of missing data was minimal for most measures. The correlations between missing-data rates and quality measure scores tend to be weak. They also tend to indicate a direct relationship between quality measure scores and complete data keeping: facilities with higher rates of missing data tend to have poorer scores (higher scores on negative measures, lower scores on positive measures). For pain-related measures, missing data rates tended to be higher for residents who had difficulty making themselves understood, indicating a specific threat to validity posed to measures requiring residents to self-report. In addition, RTI examined the impact of discharge assessments on QM scores. The inclusion of the discharge assessment in QM calculation is an important change from the MDS 2.0 to the MDS 3.0 measure (the discharge assessment was first introduced with the MDS 3.0). There has been interest in (1) the overall impact of including the discharge assessment on QM rates and (2) the completeness of the new discharge assessment items and this related impact on the QM rate. We evaluated the impact of the new discharge assessment on the QMs by examining each facility's QM score change before and after excluding discharge assessments from the set of eligible target assessments. For all measures but one, removing the discharge assessments from analyses led to a minimal reduction of the number of facilities able to report: for 14 of the 16 measures, this change represented 2 percent or fewer of the total number of facilities. Reportability actually increased after excluding discharge assessments for QM #0677 Percent of Residents Who Self-Report Moderate to Severe Pain (Long Stay) by 5.3 percentage points, owing to a high rate of missing items (and thus exclusions) on pain-related items on discharge assessments. However, the gains in reportability are substantial for QM #0678, Percent of Residents With Pressure Ulcers That Are New or Worsened (Short Stay). Discharge assessments helped 15.6 percent of facilities report that could not without the discharge assessment.

Seasonal Variation. If a QM score varies substantially from quarter to quarter in a consistent pattern over multiple years corresponding to changes in seasons, the measure's validity is suspect and likely influenced by factors outside of the nursing home's control. RTI examined variation in mean and median scores for each quality measure for each quarter in 2011 and found no evidence to suggest that seasonal variation may post a threat to validity. The widest variances in scores from quarter to quarter were seen for the two flu vaccine measures: for each, the range between the highest and lowest mean quarter scores was about 6 percentage

points. However, these results are based on just one cycle of seasons and should be considered preliminary.

Risk Adjustment. Similar to the risk-adjustment methods employed with the MDS 2.0 QMs, all but two MDS 3.0 QMs use exclusion or sample restriction; indirect standardization was used for three QMs. Exclusion or sample restriction almost always includes restrictions based on data availability and sometimes include sample exclusions related to risk factors or residents' clinical status. For the QMs that risk-adjustment models are applied in the calculation of the QM score, RTI examined the impact of the selection of covariates on QM score, changes in the impact of covariates on the QM scores across quarter, and the impact of risk-adjustment model specifications on the QM scores. The results suggest that the current risk-adjustment models show satisfactory predictive power, but using hierarchical model specifications can improve the predictive power. The selection of model specification had sizable effect on the risk adjusted QM scores and facility rank based on these QM scores.

5.1 Summary

The 16 quality measures considered in this report perform well across analytic domains. Alterations to the definitions of short- and long-stay residents and the introduction of discharge assessments in the MDS 3.0 provided modest improvements in the number of residents and facilities for which process and outcome could be measured with minimal impact on national mean measure scores. There are few specific areas that merit close monitoring for individual measures. Generally, analyses of MDS 3.0 data for the year 2011 indicate that these 16 measures have scores that vary widely enough to discriminate between facilities with different levels of quality of care and are reliable and valid.

REFERENCES

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

ACIP. (1997). Prevention of pneumococcal disease: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR. Recomm Rep. *46*(RR-8), 1-24.

Agency for Healthcare Research and Quality. (January 2009). Agency news and notes: pressure ulcers are increasing among hospital patients Retrieved from http://www.ahrq.gov/research/jan09/0109RA22.htm

AHRQ. (July 2009). On-time pressure ulcer healing project Retrieved from http://www.ahrq.gov/research/pressureulcerhealing/

Allen, R. S., Thorn, B. E., Fisher, S. E., Gerstle, J., Quarles, K., Bourgeois, M. S., Dijkstra, K., & Burgio, L. D. (2003). Prescription and dosage of analgesic medication in relation to resident behaviors in the nursing home. [Research Support, U.S. Gov't, P.H.S.]. *J Am Geriatr Soc, 51*(4), 534-538.

Amador, L. F., Al Snih, S., Markides, K. S., & Goodwin, J. S. (2006). Weight change and mortality among older Mexican Americans. *Aging Clin Exp Res, 18*(3), 196-204. doi:1458 [pii]

Amella, E. J. (2004). Feeding and hydration issues for older adults with dementia. *Nurs Clin North Am, 39*(3), 607-623. doi:10.1016/j.cnur.2004.02.014; S0029646504000350 [pii]

American Health Care Association. (2011). *Trends in Publicly Reported Nursing Facility Quality Measures*. Retrieved from http://www.ahcancal.org/research_data/trends_statistics/Documents/trends_nursing_facilities_quality_measures.pdf.

Anderson, R. L., Buckwalter, K. C., Buchanan, R. J., Maas, M. L., & Imhof, S. L. (2003). Validity and reliability of the Minimum Data Set Depression Rating Scale (MDSDRS) for older adults in nursing homes. *Age Ageing, 32*(4), 435-438.

Ang, Y. H., Au, S. Y., Yap, L. K., & Ee, C. H. (2006). Functional decline of the elderly in a nursing home. *Singapore Med J, 47*(3), 219-224.

Arling, G., Lewis, T., Kane, R. L., Mueller, C., & Flood, S. (2007). Improving quality assessment through multilevel modeling: the case of nursing home compare. *Health Serv Res, 42*(3 Pt 1), 1177-1199. doi:HESR647 [pii]; 10.1111/j.1475-6773.2006.00647.x

Bates-Jensen, B. M. (2001). Quality indicators for prevention and management of pressure ulcers in vulnerable elders. *Ann Intern Med, 135*(8 Pt 2), 744-751. doi:200110161-00014 [pii]

Bellows, N. M., & Halpin, H. A. (2008). MDS-based state Medicaid reimbursement and the ADL-decline quality indicator. *Gerontologist, 48*(3), 324-329. doi:48/3/324 [pii]

Boltz, M., Resnick, B., Capezuti, E., Shuluk, J., & Secic, M. (2012). Functional Decline in Older Adults: Can Nursing Make a Difference? *Geriatric Nursing*.

Bourdel-Marchasson, I. (2010). How to improve nutritional support in geriatric institutions. *J Am Med Dir Assoc, 11*(1), 13-20. doi:S1525-8610(09)00136-4 [pii]; 10.1016/j.jamda.2009.04.003

Brega, A., Goodrich, G., Hittle, D., Conway, K., & Levy, C. (2008). Empirical review and validation of refined pressure ulcer quality measures draft. In I. Abt Associates, (Ed.). Denver: Division of Health Care Policy and Research University of Colorado at Denver.

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

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

Capezuti, E., Evans, L., Strumpf, N., & Maislin, G. (1996). Physical restraint use and falls in nursing home residents. *J Am Geriatr Soc, 44*(6), 627-633.

Capezuti, E., Maislin, G., Strumpf, N., & Evans, L. K. (2002). Side rail use and bed-related fall outcomes among nursing home residents. *J Am Geriatr Soc, 50*(1), 90-96. doi:50013 [pii]

Carter, M. W., & Porell, F. W. (2006). Nursing home performance on select publicly reported quality indicators and resident risk of hospitalization: grappling with policy implications. *J Aging Soc Policy, 18*(1), 17-39. doi:10.1300/J031v18n01_02

Castle, N., Degenholtz, H., & Engberg, J. (2005). State variability in indicators of quality of care in nursing facilities. *J Gerontol., 60A*(9), 1173-1179.

Castle, N. G. (2002). Nursing homes with persistent deficiency citations for physical restraint use. *Med Care, 40*(10), 868-878. doi:10.1097/01.MLR.0000027357.74348.20

Castle, N. G., & Mor, V. (1998). Physical restraints in nursing homes: a review of the literature since the Nursing Home Reform Act of 1987. *Med Care Res Rev, 55*(2), 139-170; discussion 171-136.

Centers for Disease Control and Prevention. State specific influenza vaccination coverage among adults – United States, 2006-2007 influenza season. *MMWR, 57*(38), 1033-1039.

Centers for Disease Control and Prevention. (1997). *Pneumococcal polysaccharide vaccine. What you need to know*. Atlanta, GA: CDC.

Centers for Disease Control and Prevention. (2005). Data & Statistics (WISQARS™): Cost of Injury Reports Retrieved from http://wisqars.cdc.gov:8080/costT/

Centers for Disease Control and Prevention. (2008). Influenza E-brief: 2008-2009 flu facts for policymakers. (September 2008), 1-2. Retrieved from http://www.cdc.gov/washington/pdf/flu_newsletter.pdf website: http://www.cdc.gov/washington/pdf/flu_newsletter.pdf

Centers for Disease Control and Prevention. (2009). 10 Leading Causes of Death by Age Group, United States – 2009 Retrieved September 17, 2012, from http://www.cdc.gov/Injury/wisqars/pdf/10LCD-Age-Grp-US-2009-a.pdf

Centers for Disease Control and Prevention. (2010). Prevalence and Trends Data: Immunization--2010: Adults Aged 65+ Who Have Ever Had a Pneumonia Vaccination Retrieved September 18, 2012, from http://apps.nccd.cdc.gov/BRFSS/list.asp?cat=IM&yr=2010&qkey=4408&state=All

Centers for Medicare & Medicaid Services. MDS Quality Measure/Indicator Report Retrieved from http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MDSPubQIandResRep/qmreport.html

Centers for Medicare & Medicaid Services. (2003). *MDS quality measure/indicator report*. Retrieved from http://www.cms.hhs.gov/MDSPubQIandResRep/02_qmreport.asp#TopOfPage.

Centers for Medicare & Medicaid Services. (2006). Appendix PP, Guidance to Surveyors for Long Term Care Facilities, rev. 15. F315 urinary incontinence *State Operations Manual*. Baltimore, MD.

Centers for Medicare & Medicaid Services. (2007). *Health Care Financing Review. Statistical Supplement 293*. Baltimore, MD: Centers for Medicare & Medicaid Services.

Centers for Medicare & Medicaid Services. (2008). *MDS quality measure/indicator report*. Retrieved from http://www.cms.gov/apps/mds/default.asp

Centers for Medicare & Medicaid Services. (2009). MDS quality measure/indicator report Retrieved from http://www.cms.hhs.gov/MDSPubQIandResRep/02_qmreport.asp#TopOfPage

Centers for Medicare & Medicaid Services. (2010a). CMS Nursing Home Data Compendium, 2010 Retrieved September 17, 2012, from https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandComplianc/downloads/nursinghomedatacompendium_508.pdf

Centers for Medicare & Medicaid Services. (2010b). HHS/CMS Resident Assessment Instrument (RAI) Version 3.0 Manual Retrieved from http://www.cms.gov/NursingHomeQualityInits/45_NHQIMDS30TrainingMaterials.asp#TopOfPage

Centers for Medicare & Medicaid Services. (2010c, July/September). MDS Quality Measure/Indicator Report Infection Control Retrieved from http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MDSPubQIandResRep/qmreport.html

Centers for Medicare & Medicaid Services. (2011, May). MDS quality measure/indicator report Retrieved from http://www.cms.hhs.gov/MDSPubQIandResRep/02_qmreport.asp#TopOfPage

Centers for Medicare & Medicaid Services. (2012). MDS Quality Measure/Indicator Report Retrieved from http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MDSPubQIandResRep/qmreport.html

Chen, J. H., Chan, D. C., Kiely, D. K., Morris, J. N., & Mitchell, S. L. (2007). Terminal trajectories of functional decline in the long-term care setting. *J Gerontol A Biol Sci Med Sci, 62*(5), 531-536. doi:62/5/531 [pii]

Chodosh, J., Edelen, M. O., Buchanan, J. L., Yosef, J. A., Ouslander, J. G., Berlowitz, D. R., Streim, J. E., & Saliba, D. (2008). Nursing home assessment of cognitive impairment: development and testing of a brief instrument of mental status. *J Am Geriatr Soc, 56*(11), 2069-2075. doi:10.1111/j.1532-5415.2008.01944.x

Closs, S. J., Cash, K., Barr, B., & Briggs, M. (2005). Cues for the identification of pain in nursing home residents. [Research Support, Non-U.S. Gov't]. *Int J Nurs Stud, 42*(1), 3-12. doi:10.1016/j.ijnurstu.2004.05.012

Colorado Foundation for Medical Care. (2007). Environmental scan: review of the literature, clinical guidelines, and other sources of information pertinent to the CMS publicly reported nursing home quality measures. Englewood, CO: Colorado Foundation for Medical Care.

Colorado Foundation for Medical Care. (2011, February 18). Quarterly Trends of Campaign Nursing Home Group Means: High-Risk Pressure Ulcers Retrieved September 19, 2012, from http://www.nhqualitycampaign.org/files/reports/results/q3-2010/PressureUlcers_ClinicalPerformanceTrends_2010Q3.pdf

Covinsky, K. E., Martin, G. E., Beyth, R. J., Justice, A. C., Sehgal, A. R., & Landefeld, C. S. (1999). The relationship between clinical assessments of nutritional status and adverse outcomes in older hospitalized medical patients. *J Am Geriatr Soc, 47*(5), 532-538.

Datto, C. J., Oslin, D. W., Streim, J. E., Scheinthal, S. M., DiFilippo, S., & Katz, I. R. (2002). Pharmacologic treatment of depression in nursing home residents: a mental health services perspective. *J Geriatr Psychiatry Neurol, 15*(3), 141-146.

Degenholtz, H. B., Rosen, J., Castle, N., Mittal, V., & Liu, D. (2008). The association between changes in health status and nursing home resident quality of life. *Gerontologist, 48*(5), 584-592. doi:48/5/584 [pii]

Duncan, K. D. (2007). Preventing pressure ulcers: The goal is zero. *Jt Comm J Qual Patient Saf., 33*(10), 605-610.

Ensrud, K. E., Ewing, S. K., Stone, K. L., Cauley, J. A., Bowman, P. J., & Cummings, S. R. (2003). Intentional and unintentional weight loss increase bone loss and hip fracture risk in older women. *J Am Geriatr Soc, 51*(12), 1740-1747. doi:51558 [pii]

Evans, L. K., Strumpf, N. E., Allen-Taylor, S. L., Capezuti, E., Maislin, G., & Jacobsen, B. (1997). A clinical trial to reduce restraints in nursing homes. *J Am Geriatr Soc, 45*(6), 675-681.

Feinsod, F. M., Levenson, S. A., Rapp, K., Rapp, M. P., Beechinor, E., & Liebmann, L. (2004). Dehydration in frail, older residents in long-term care facilities. *J Am Med Dir Assoc, 5*(2 Suppl), S35-41. doi:10.1097/01.JAM.0000036697.72589.A8; 01.JAM.0000036697.72589.A8 [pii]

Feinsod, F. M., Moore, M., & Levenson, S. (1997). Eliminating full-length bed rails from long term care facilities. *Nurs Home Med, 5*, 257-263.

Fennell, M. L., Feng, Z., Clark, M. A., & Mor, V. (2010). Elderly hispanics more likely to reside in poor-quality nursing homes. [Comparative Study

Research Support, N.I.H., Extramural]. *Health Aff (Millwood), 29*(1), 65-73. doi:10.1377/hlthaff.2009.0003

Ferrell, B. A., Ferrell, B. R., & Osterweil, D. (1990). Pain in the nursing home. *J Am Geriatr Soc, 38*(4), 409-414.

Finnegan, T., Austin, T., & Cape, R. (1985). 12-month fever surveillance study in a veterans' long-stay institution. *J Am Geriatr Soc., 33*, 590-594.

Fry, A. M., Shay, D. K., Holman, R. C., Curns, A. T., & Anderson, L. J. (2005). Trends in hospitalizations for pneumonia among persons aged 65 years or older in the United States, 1988-2002. *JAMA, 294*(21), 2712-2719. doi:294/21/2712 [pii]; 10.1001/jama.294.21.2712

Gammack, J. (2003). Use and management of chronic urinary catheters in the nursing home: Much controversy, little consensus. *J Am Med Dir Assoc., 4*(2 Supp), S52-59.

Gammack, J. K. (2003). Use and management of chronic urinary catheters in long-term care: much controversy, little consensus. *J Am Med Dir Assoc, 4*(2 Suppl), S52-59.

Gastmans, C., & Milisen, K. (2006). Use of physical restraint in nursing homes: clinical-ethical considerations. *J Med Ethics, 32*(3), 148-152. doi:32/3/148 [pii]; 10.1136/jme.2005.012708

Georgiou, A., Potter, J., Brocklehurst, J. C., Lowe, D., & Pearson, M. (2001). Measuring the quality of urinary continence care in long-term care facilities: an analysis of outcome indicators. *Age Ageing, 30*(1), 63-66.

Gorina, Y., Kelly, T., Lubitz, J., & Hines, Z. (2008). Trends in influenza and pneumonia among older persons in the United States. *Aging Trends*(8), 1-11.

Gould, C., Umscheid, C., Agarwal, R., Kuntz, G., Pegues, D., & the Healthcare Infection Control Practices Advisory Committee. (2009). Guideline for prevention of catheter-associated urinary tract infections 2009 Retrieved from http://www.cdc.gov/ncidod/dhqp/pdf/guidelines/CAUTI_Guideline2009final.pdf

Grabowski, D. C. (2004). The admission of blacks to high-deficiency nursing homes. *Med Care, 42*(5), 456-464. doi:00005650-200405000-00009 [pii]

Grant, M., Ferrell, B. R., Rivera, L. M., & Lee, J. (1995). Unscheduled readmissions for uncontrolled symptoms. A health care challenge for nurses. *Nurs Clin North Am, 30*(4), 673-682.

Greenberg, P. E., Kessler, R. C., Birnbaum, H. G., Leong, S. A., Lowe, S. W., Berglund, P. A., & Corey-Lisle, P. K. (2003). The economic burden of depression in the United States: how did it change between 1990 and 2000? *J Clin Psychiatry, 64*(12), 1465-1475.

Gulpers, M. J., Bleijlevens, M. H., Ambergen, T., Capezuti, E., van Rossum, E., & Hamers, J. P. (2011). Belt restraint reduction in nursing homes: effects of a multicomponent intervention program. *Journal of the American Geriatrics Society, 59*(11), 2029-2036.

Gumieiro, D. N., Rafacho, B. P., Gradella, L. M., Azevedo, P. S., Gaspardo, D., Zornoff, L. A., Pereira, G. J., Paiva, S. A., & Minicucci, M. F. (2012). Handgrip strength predicts pressure ulcers in patients with hip fractures. *Nutrition, 28*(9), 874-878. doi:10.1016/j.nut.2011.11.010

Hanson, L. (1997). Can clinical interventions change care at the end of life? Annals of Internal Medicine. *126, 5*, 381-388.

Heston, L. L., Garrard, J., Makris, L., Kane, R. L., Cooper, S., Dunham, T., & Zelterman, D. (1992). Inadequate treatment of depressed nursing home elderly. *J Am Geriatr Soc, 40*(11), 1117-1122.

Horgas, A. L., & Tsai, P. F. (1998). Analgesic drug prescription and use in cognitively impaired nursing home residents. *Nurs Res, 47*(4), 235-242.

Horn, S. D., Sharkey, S. S., Hudak, S., Smout, R. J., Quinn, C. C., Yody, B., & Fleshner, I. (2010). Beyond CMS quality measure adjustments: identifying key resident and nursing home facility factors associated with quality measures. *Journal of the American Medical Directors Association, 11*(7), 500-505.

Howard, D. L., Sloane, P. D., Zimmerman, S., Eckert, J. K., Walsh, J. F., Buie, V. C., Taylor, P. J., & Koch, G. G. (2002). Distribution of African Americans in residential care/assisted living and nursing homes: more evidence of racial disparity? *Am J Public Health, 92*(8), 1272-1277.

Huffman, G. B. (2002). Evaluating and treating unintentional weight loss in the elderly. [Review]. *Am Fam Physician, 65*(4), 640-650.

Hurd, D., Moore, T., Radley, D., & Williams, C. (2010). Pressure Ulcer Prevalence and Incidence across Post-acute Care Settings. Home Health Quality Measures & Data Analysis Project, Report of Findings, prepared for CMS/OCSQ, Baltimore, MD, under Contract No. 500-2005-000181 TO 0002.

Hurd, D., T., M., Radley, D., & Williams, C. (2010). Pressure ulcer prevalence and incidence across post-acute care settings. Report of Findings. *Home Health Quality Measures & Data Analysis Project*. Baltimore, MD: Centers for Medicare & Medicaid Services.

Institute for Healthcare Improvement. (2007). Relieve the pressure and reduce harm Retrieved from http://www.ihi.org/knowledge/Pages/ImprovementStories/RelievethePressureandReduceHarm.aspx

Institute for Healthcare Improvement. (2011, August 18). Relieve the Pressure and Reduce Harm Retrieved from http://www.ihi.org/knowledge/Pages/ImprovementStories/RelievethePressureandReduceHarm.aspx

Institute of Medicine. (1986). *Improving the quality of care in nursing homes*. Washington, D.C.: National Academies Press.

Iowa Foundation for Medical Care, & University of Michigan Stepwise Systems CareTrack. (2005). Staff Time and Resource Intensity Verification Project. Prepared for Centers for Medicare & Medicaid Services under Systems Contract Number: 500-02-0030\0002.

Jackson, M., Fierer, J., Barrett-Conner, E., & et al. (1992). Intensive surveillance for infections in a three year study of nursing home patients. *Am J Epidemiol., 135*, 685-696.

Kaldy, J. (2000). Clinical issues in weight loss and dehydration. *J Am Med Dir Assoc, 1*(6 Suppl), S35-36.

Kamel, H. K. (2004). Managing urinary tract infections in the nursing home: Myths, mysteries and realities. *Int J Geriatr Gerontol, 1*(2).

Katz, P., Beam, T. J., Brand, F., & et al. (1990). Antibiotic use in the nursing home. Physician practice patterns. *Arch Int Med., 150*, 1465-1468.

Kayser-Jones, J., & Pengilly, K. (1999). Dysphagia among nursing home residents. *Geriatr Nurs, 20*(2), 77-82; quiz 84. doi:S0197-4572(99)00024-5 [pii]; 10.1053/gn.1999.v20.97011

Keller, H. H., & Ostbye, T. (2005). Body Mass Index (BMI), BMI change and mortality in community-dwelling seniors without dementia. *J Nutr Health Aging, 9*(5), 316-320.

Kiely, D. K., & Flacker, J. M. (2000). Resident characteristics associated with mortality in long-term care nursing homes: is there a gender difference? *J Am Med Dir Assoc, 1*(1), 8-13.

Koczy, P., Becker, C., Rapp, K., Klie, T., Beische, D., Buchele, G., Kleiner, A., Guerra, V., Rissmann, U., Kurrle, S., & Bredthauer, D. (2011). Effectiveness of a multifactorial intervention to reduce physical restraints in nursing home residents. *Journal of the American Geriatrics Society, 59*(2), 333-339.

Kusserow, R. (1991). Minimizing restraints in nursing homes: a guide to action. Baltimore, MD: Centers for Medicare & Medicaid Services.

Landefeld, C. S., Bowers, B. J., Feld, A. D., Hartmann, K. E., Hoffman, E., Ingber, M. J., King, J. T., Jr., McDougal, W. S., Nelson, H., Orav, E. J., Pignone, M., Richardson, L. H., Rohrbaugh, R. M., Siebens, H. C., & Trock, B. J. (2008). National Institutes of Health state-of-the-science conference statement: prevention of fecal and urinary incontinence in adults. *Ann Intern Med, 148*(6), 449-458. doi:0000605-200803180-00210 [pii]

Langlois, J. A., Harris, T., Looker, A. C., & Madans, J. (1996). Weight change between age 50 years and old age is associated with risk of hip fracture in white women aged 67 years and older. *Arch Intern Med, 156*(9), 989-994.

Langlois, J. A., Mussolino, M. E., Visser, M., Looker, A. C., Harris, T., & Madans, J. (2001). Weight loss from maximum body weight among middle-aged and older white women and the risk of hip fracture: the NHANES I epidemiologic follow-up study. *Osteoporos Int, 12*(9), 763-768.

Lapane, K. L., & Resnik, L. (2005). Obesity in nursing homes: an escalating problem. *Journal of the American Geriatrics Society, 53*(8), 1386-1391.

Lee, Y., Thrupp, L., Friis, H., & et al. (1992). Nosocomial infection and antibiotic utilization in geriatric patients: a pilot prospective surveillance program in skilled nursing facilities. *Gerontology, 38*, 223-232.

Lekan-Rutledge, D., & Colling, J. (2003). Urinary incontinence in the frail elderly: even when it's too late to prevent a problem, you can still slow its progress. *Am J Nurs, Suppl*, 36-46.

Leone, A. F., Standoli, F., & Hirth, V. (2009). Implementing a pain management program in a long-term care facility using a quality improvement approach. *J Am Med Dir Assoc, 10*(1), 67-73. doi:10.1016/j.jamda.2008.08.003

Leung, F. W., & Schnelle, J. F. (2008). Urinary and fecal incontinence in nursing home residents. *Gastroenterol Clin North Am, 37*(3), 697-707, x. doi:10.1016/j.gtc.2008.06.005

Luo, H., Lin, M., & Castle, N. (2011). Physical restraint use and falls in nursing homes: a comparison between residents with and without dementia. *American Journal of Alzheimer's Disease & Other Dementias, 26*(1), 44-50.

Lynn, J., West, J., Hausmann, S., Gifford, D., Nelson, R., McGann, P., Bergstrom, N., & Ryan, J. A. (2007). Collaborative clinical quality improvement for pressure ulcers in nursing homes. *Journal of the American Geriatrics Society, 55*(10), 1663-1669.

MacLean, D. S. (2003). Preventing & managing pressure sores. *Caring for the Ages, 4*(3), 34-35.

Magaziner, J., Lydick, E., Hawkes, W., Fox, K. M., Zimmerman, S. I., Epstein, R. S., & Hebel, J. R. (1997). Excess mortality attributable to hip fracture in white women aged 70 years and older. *Am J Public Health, 87*(10), 1630-1636.

Magaziner, J., Tenney, J., Deforge, B., & et al. (1991). Prevalence and Characteristics of Nursing Home-Acquired Infections in the Aged. *J Am Geriatr Soc., 39*, 1071-1078.

McConnell, E. S., Pieper, C. F., Sloane, R. J., & Branch, L. G. (2002). Effects of cognitive performance on change in physical function in long-stay nursing home residents. *J Gerontol A Biol Sci Med Sci, 57*(12), M778-784.

McCurren, C. (2002). Assessment for depression among nursing home elders: evaluation of the MDS mood assessment. *Geriatr Nurs, 23*(2), 103-108. doi:S0197457202420381 [pii]

MedQIC. (2002). Quality Measures Management Information System (QMIS) Retrieved from https://www.qualitynet.org/qmis/measureDetailView.htm?measureId=10176&viewType=0

MedQIC. (2004). Nursing Home Quality Initiative. Fast facts: urinary catheters overview Retrieved from http://www.njha.com/qualityinstitute/pdf/118200631453PM85.pdf

Milenkovic, M., Russo, C. A., & Elixhauser, A. (2006). Hospital stays for influenza, 2004. *H-CUP Healthcare cost and utilization project*. Rockville, MD: Agency for Healthcare Research and Quality.

Miller, S. C., Papandonatos, G., Fennell, M., & Mor, V. (2006). Facility and county effects on racial differences in nursing home quality indicators. [Research Support, N.I.H., Extramural]. *Soc Sci Med, 63*(12), 3046-3059. doi:10.1016/j.socscimed.2006.08.003

Mohler, R., Richter, T., Kopke, S., & Meyer, G. (2011). Interventions for preventing and reducing the use of physical restraints in long-term geriatric care. *Cochrane Database Syst Rev*(2), CD007546. doi:10.1002/14651858.CD007546.pub2

Mor, V., Gruneir, A., Feng, Z., Grabowski, D. C., Intrator, O., & Zinn, J. (2011). The effect of state policies on nursing home resident outcomes. *Journal of the American Geriatrics Society, 59*(1), 3-9.

Mor, V., Zinn, J., Angelelli, J., Teno, J. M., & Miller, S. C. (2004). Driven to tiers: socioeconomic and racial disparities in the quality of nursing home care. *Milbank Q, 82*(2), 227-256. doi:10.1111/j.0887-378X.2004.00309.x; MILQ309 [pii]

Morley, J. E. (2007). Weight loss in the nursing home. *J Am Med Dir Assoc, 8*(4), 201-204. doi:S1525-8610(07)00131-4 [pii]; 10.1016/j.jamda.2007.02.003

Murden, R. A., & Ainslie, N. K. (1994). Recent weight loss is related to short-term mortality in nursing homes. *J Gen Intern Med, 9*(11), 648-650.

National Nursing Home Survey. (2004). Number and Percent Distribution of Nursing Home Residents by Pneumococcal Vaccination Status According to Age, Sex, and Race: United States, 2004 Retrieved from http://www.cdc.gov/nchs/nnhs/resident_tables_estimates.htm

National Pressure Ulcer Advisory Panel Board of Directors. (2001). Pressure ulcers in America: prevalence, incidence, and implications for the future. An executive summary of the National Pressure Ulcer Advisory Panel Monograph. *Adv Skin Wound Care., 14*(4), 208-215.

National Quality Forum. (2008). National voluntary consensus standards for influenza and pneumococcal immunizations Retrieved from http://www.qualityforum.org/Publications/2008/12/National_Voluntary_Consensus_Standards_for_Influenza_and_Pneumococcal_Immunizations.aspx

National Quality Forum. (2010). Review of Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties, Draft Report for Review and Comment Retrieved from http://www.qualityforum.org/Measuring_Performance/Improving_NQF_Process/Measure_Testing_Task_Force.aspx

Newman, D. K., Fader, M., & Bliss, D. Z. (2004). Managing incontinence using technology, devices, and products: directions for research. *Nurs Res, 53*(6 Suppl), S42-48. doi:00006199-200411006-00007 [pii]

Newman, D. K., & Palmer, M. H. (1999). Incontinence and PPS: a new era. *Ostomy Wound Manage, 45*(12), 32-44, 46, 48-39.

Nicolle, L. (1993). Urinary tract infections in long-term care facilities. Infect Control Hosp. *Epidemiol., 14*, 220-225.

Nicolle, L., McIntyre, M., Zacharies, H., & et al. (1984). Twelve month surveillance of infections in institutionalized elderly men. *J Am Geriatr Soc., 32*, 513-519.

Nicolle, L., & SHEA Long-Term-Care Committee. (2001). Urinary tract infections in long-term care facilities. Infect Control Hosp. *Epidemiol., 22*, 167-175.

Offermans, M., Du Moulin, M., Hamers, J., Dassen, T., & Halfens, R. (2009). Prevalence of urinary incontinence and associated risk factors in nursing home residents: a systematic review. *Neurourol Urodyn., 28*(4), 288-294.

Ouslander, J. G., & Schnelle, J. F. (1995). Incontinence in the nursing home. *Ann Intern Med, 122*(6), 438-449.

Park-Lee, E., & Caffrey, C. (2009). Pressure ulcers among nursing home residents: United States, 2004 (NCHS Data Brief No. 14). *National Center for Health Statistics*. Retrieved from http://www.cdc.gov/nchs/data/databriefs/db14.htm

Parker, K., & Miles, S. H. (1997). Deaths caused by bedrails. *J Am Geriatr Soc, 45*(7), 797-802.

Parmelee, P. A., Katz, I. R., & Lawton, M. P. (1989). Depression among institutionalized aged: assessment and prevalence estimation. *J Gerontol, 44*(1), M22-29.
Parmelee, P. A., Smith, B., & Katz, I. R. (1993). Pain complaints and cognitive status among elderly institution residents. *J Am Geriatr Soc, 41*(5), 517-522.
Pauly, L., Stehle, P., & Volkert, D. (2007). Nutritional situation of elderly nursing home residents. *Z Gerontol Geriatr, 40*(1), 3-12. doi:10.1007/s00391-007-0430-x
Peek, C. W., Coward, R. T., Henretta, J. C., Duncan, R. P., & Dougherty, M. C. (1997). Differences by race in the decline of health over time. *J Gerontol B Psychol Sci Soc Sci, 52*(6), S336-344.
Phillips, C. D., Hawes, C., & Fries, B. E. (1993). Reducing the use of physical restraints in nursing homes: will it increase costs? *Am J Public Health, 83*(3), 342-348.
Phillips, C. D., Shen, R., Chen, M., & Sherman, M. (2007). Evaluating nursing home performance indicators: an illustration exploring the impact of facilities on ADL change. *Gerontologist, 47*(5), 683-689. doi:47/5/683 [pii]
Quality Measures Management Information System (QMIS). Measure details: Measure 10177 Retrieved from https://www.qualitynet.org/qmis/measureDetailView.htm?measureId=10177&viewType=1
Quality Measures Management Information System (QMIS). (2002). Measure details Retrieved from https://www.qualitynet.org/qmis/measureDetailView.htm?measureId=10176&viewType=0
Rogers, M. A., Fries, B. E., Kaufman, S. R., Mody, L., McMahon, L. F., Jr., & Saint, S. (2008). Mobility and other predictors of hospitalization for urinary tract infection: a retrospective cohort study. *BMC Geriatrics, 8*(31).
Rogers, M. A., Mody, L., Kaufman, S. R., Fries, B. E., McMahon, L. F., Jr., & Saint, S. (2008). Use of urinary collection devices in skilled nursing facilities in five states. *J Am Geriatr Soc, 56*(5), 854-861. doi:10.1111/j.1532-5415.2008.01675.x
Rubenstein, L. Z., Josephson, K. R., & Robbins, A. S. (1994). Falls in the nursing home. *Ann Intern Med, 121*(6), 442-451.
Russo, C. A., Steiner, C., & Spector, W. (2006). Hospitalizations Related to Pressure Ulcers Among Adults 18 Years and Older, 2006: Statistical Brief #64. doi:NBK54557 [bookaccession]
Russo, C. A., Steiner, C., & Spector, W. (2008). Hospitalizations related to pressure ulcers among adults 18 years and older, 2006 (Healthcare Cost and Utilization Project Statistical Brief No. 64) Retrieved from http://www.hcup-us.ahrq.gov/reports/statbriefs/sb64.pdf
Ryan, C., Bryant, E., Eleazer, P., Rhodes, A., & Guest, K. (1995). Unintentional weight loss in long-term care: predictor of mortality in the elderly. *South Med J, 88*(7), 721-724.
Sachs, G. A., Shega, J. W., & Cox-Hayley, D. (2004). Barriers to excellent end-of-life care for patients with dementia. *J Gen Intern Med, 19*(10), 1057-1063. doi:JGI30329 [pii]; 10.1111/j.1525-1497.2004.30329.x
Saint, S., Kauman, S., Robers, M., & et al. (2006). Risk factors for nosocomial urinary tract-related bacteremia: A case control study. *Am J Infect Control., 34*(7), 401-407.
Saliba, D., & Buchanan, J. (2008). Development and validation of a revised nursing home assessment tool: MDS 3.0. Santa Monica, CA: Rand Health Corporation.
Scherder, E. J., & Bouma, A. (2000). Visual analogue scales for pain assessment in Alzheimer's disease. *Gerontology, 46*(1), 47-53. doi:22133 [pii]

Schnelle, J. F., Leung, F. W., Rao, S. S., Beuscher, L., Keeler, E., Clift, J. W., & Simmons, S. (2010). A controlled trial of an intervention to improve urinary and fecal incontinence and constipation. *J Am Geriatr Soc, 58*(8), 1504-1511. doi:10.1111/j.1532-5415.2010.02978.x

Sengstaken, E. A., & King, S. A. (1993). The problems of pain and its detection among geriatric nursing home residents. *J Am Geriatr Soc, 41*(5), 541-544.

Simmons, S. F., Garcia, E. T., Cadogan, M. P., Al-Samarrai, N. R., Levy-Storms, L. F., Osterweil, D., & Schnelle, J. F. (2003). The minimum data set weight-loss quality indicator: does it reflect differences in care processes related to weight loss? *J Am Geriatr Soc, 51*(10), 1410-1418. doi:51459 [pii]

Sloane, P. D., Ivey, J., Helton, M., Barrick, A. L., & Cerna, A. (2008). Nutritional issues in long-term care. *J Am Med Dir Assoc, 9*(7), 476-485. doi:S1525-8610(08)00103-5 [pii]; 10.1016/j.jamda.2008.03.005

Smith, D. B., Feng, Z., Fennell, M. L., Zinn, J. S., & Mor, V. (2007). Separate and unequal: racial segregation and disparities in quality across U.S. nursing homes. *Health Aff (Millwood), 26*(5), 1448-1458. doi:26/5/1448 [pii]; 10.1377/hlthaff.26.5.1448

Smith, P. A. (2006). Nutrition, hydration, and dysphagia in long-term care: Differing opinions on the effects of aspiration. *J Am Med Dir Assoc, 7*(9), 545-549. doi:S1525-8610(06)00178-2 [pii]; 10.1016/j.jamda.2006.03.008

Snowden, M., Sato, K., & Roy-Byrne, P. (2003). Assessment and treatment of nursing home residents with depression or behavioral symptoms associated with dementia: a review of the literature. *J Am Geriatr Soc, 51*(9), 1305-1317. doi:51417 [pii]

Sorensen, S. V., de Lissovoy, G., Kunaprayoon, D., Resnick, B., Rupnow, M. F., & Studenski, S. (2006). A taxonomy and economic consequences of nursing home falls. *Drugs Aging, 23*(3), 251-262. doi:2337 [pii]

Stevens, J. A., Thomas, K., Teh, L., & Greenspan, A. I. (2009). Unintentional fall injuries associated with walkers and canes in older adults treated in U.S. emergency departments. *J Am Geriatr Soc, 57*(8), 1464-1469. doi:10.1111/j.1532-5415.2009.02365.x

Strausbaugh, L., & Joseph, C. (2000). The burden of infection in long-term care. *Infect Control Hosp Epidemiol., 21*(10), 674-679.

Sullivan-Marx, E. M. (2001). Achieving restraint-free care of acutely confused older adults. *J Gerontol Nurs, 27*(4), 56-61.

Sullivan-Marx, E. M., Strumpf, N. E., Evans, L. K., Baumgarten, M., & Maislin, G. (1999). Initiation of physical restraint in nursing home residents following restraint reduction efforts. *Res Nurs Health, 22*(5), 369-379. doi:10.1002/(SICI)1098-240X(199910)22:5<369::AID-NUR3>3.0.CO;2-G [pii]

Sullivan, D. H., Morley, J. E., Johnson, L. E., Barber, A., Olson, J. S., Stevens, M. R., Yamashita, B. D., Reinhart, S. P., Trotter, J. P., & Olave, X. E. (2002). The GAIN (Geriatric Anorexia Nutrition) registry: the impact of appetite and weight on mortality in a long-term care population. *J Nutr Health Aging, 6*(4), 275-281.

Teno, J. M., Gozalo, P., Mitchell, S. L., Kuo, S., Fulton, A. T., & Mor, V. (2012). Feeding tubes and the prevention or healing of pressure ulcers. [Research Support, N.I.H., Extramural]. *Arch Intern Med, 172*(9), 697-701. doi:10.1001/archinternmed.2012.1200

Teresi, J., Abrams, R., Holmes, D., Ramirez, M., & Eimicke, J. (2001). Prevalence of depression and depression recognition in nursing homes. *Soc Psychiatry Psychiatr Epidemiol, 36*(12), 613-620.

Thomas, D. R., Ashmen, W., Morley, J. E., & Evans, W. J. (2000). Nutritional management in long-term care: development of a clinical guideline. Council for Nutritional Strategies in Long-Term Care. *J Gerontol A Biol Sci Med Sci, 55*(12), M725-734.

Tippet, A. W. (2009). Reducing the incidence of pressure ulcers in nursing home residents: a prospective 6-year evaluation. *Ostomy Wound Manage, 55*(11), 52-58.

Tsan, L., Davis, C., Langberg, R., Hojlo, C., Pierce, J., Miller, M., Gaynes, R., Gibert, C., Montgomery, O., Bradley, S., Richards, C., Danko, L., & Roselle, G. (2008). Prevalence of nursing home-associated infections in the Department of Veterans Affairs nursing home care units. *American Journal of Infection Control, 36*(3), 173-179.

U.S. Department of Health and Human Services. (2000). Health People 2010 Retrieved from http://www.health.gov/healthypeople

U.S. Department of Health and Human Services. (2012). Healthy People 2020 Retrieved September 13, 2012, from http://www.healthypeople.gov/2020/default.aspx

United States Senate Special Committee on Aging. (2010) Retrieved January 21, from http://aging.senate.gov/record.cfm?id=319434

Unutzer, J., Schoenbaum, M., Katon, W. J., Fan, M. Y., Pincus, H. A., Hogan, D., & Taylor, J. (2009). Healthcare costs associated with depression in medically Ill fee-for-service medicare participants. [Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't]. *J Am Geriatr Soc, 57*(3), 506-510. doi:10.1111/j.1532-5415.2008.02134.x

Van Den Bos, J., Rustagi, K., Gray, T., Halford, M., Ziemkiewicz, E., & Shreve, J. (2011). The $17.1 billion problem: the annual cost of measurable medical errors. [Research Support, Non-U.S. Gov't]. *Health Aff (Millwood), 30*(4), 596-603. doi:10.1377/hlthaff.2011.0084

Vu, M. Q., Weintraub, N., & Rubenstein, L. Z. (2004). Falls in the nursing home: are they preventable? *J Am Med Dir Assoc, 5*(6), 401-406. doi:S1525-8610(04)70009-2 [pii]; 10.1097/01.JAM.0000144553.45330.AD

Wang, J., Chang, L. H., Eberly, L. E., Virnig, B. A., & Kane, R. L. (2010). Cognition moderates the relationship between facility characteristics, personal impairments, and nursing home residents' activities of daily living. *J Am Geriatr Soc., 58*(12), 2275-2283.

Wedick, N. M., Barrett-Connor, E., Knoke, J. D., & Wingard, D. L. (2002). The relationship between weight loss and all-cause mortality in older men and women with and without diabetes mellitus: the Rancho Bernardo study. *J Am Geriatr Soc, 50*(11), 1810-1815. doi:jgs50509 [pii]

Weiner, D. K., & Rudy, T. E. (2002). Attitudinal barriers to effective treatment of persistent pain in nursing home residents. *J Am Geriatr Soc, 50*(12), 2035-2040. doi:jgs50618 [pii]

Weintraub, D., Datto, C. J., Streim, J. E., & Katz, I. R. (2002). Second-generation issues in the management of depression in nursing homes. *J Am Geriatr Soc, 50*(12), 2100-2101; author reply 2101. doi:jgs50632 [pii]

Wiener, J., Freiman, B. R., & Osterweil, D. (2007). Nursing home care quality twenty years after the Omnibus Budget Reconciliation Act of 1987. Washington, D.C.: RTI International.

Williams, C. C., & Finch, C. E. (1997). Physical restraint: not fit for woman, man, or beast. *J Am Geriatr Soc, 45*(6), 773-775.

Wrede-Seaman, L. D. (2001). Treatment options to manage pain at the end of life. *Am J Hosp Palliat Care, 18*(2), 89-101.

Wu, N., Miller, S. C., Lapane, K., & Gozalo, P. (2003). The problem of assessment bias when measuring the hospice effect on nursing home residents' pain. *J Pain Symptom Manage, 26*(5), 998-1009. doi:S0885392403003282 [pii]

Yardley, L., & Smith, H. (2002). A prospective study of the relationship between feared consequences of falling and avoidance of activity in community-living older people. *Gerontologist, 42*(1), 17-23.

Zimmer, J., Bentley, D., Valenti, W., & et al. (1986). Systemic antibiotic use in nursing homes. A quality assessment. *J Am Geriatr Soc., 34*, 703-710.