

NATIONAL QUALITY FORUM

Measure Submission and Evaluation Worksheet 5.0

This form contains the information submitted by measure developers/stewards, organized according to NQF's measure evaluation criteria and process. The evaluation criteria, evaluation guidance documents, and a blank online submission form are available on the [submitting standards web page](#).

NQF #: 0421 NQF Project: Population Health: Prevention Project
(for Endorsement Maintenance Review) Original Endorsement Date: Jul 31, 2008 Most Recent Endorsement Date: Jul 31, 2008 Last Updated Date: May 08, 2012
BRIEF MEASURE INFORMATION
De.1 Measure Title: Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up
Co.1.1 Measure Steward: Centers for Medicare and Medicaid Services
De.2 Brief Description of Measure: Percentage of patients aged 18 years and older with a calculated BMI in the past six months or during the current visit documented in the medical record AND if the most recent BMI is outside of normal parameters, a follow-up plan is documented Normal Parameters: Age 65 years and older BMI > = to 23 and <30 Age 18 – 64 years BMI > = to 18.5 and <25
2a1.1 Numerator Statement: ALL MEASURE SPECIFICATION DETAILS REFERENCE THE 2012 PHYSICIAN QUALITY REPORTING SYSTEM MEASURE SPECIFICATION. Patients with BMI calculated within the past six months or during the current visit and a follow-up plan documented if the BMI is outside of parameters
2a1.4 Denominator Statement: ALL MEASURE SPECIFICATION DETAILS REFERENCE THE 2012 PHYSICIAN QUALITY REPORTING SYSTEM MEASURE SPECIFICATION. All patients aged 18 years and older on date of encounter seen during the 12 month reporting period with one or more denominator CPT or HCPCS encounter codes reported on the Medicare Part B Claims submission for the encounter along with one of the 6 numerator HCPCS clinical quality codes. All discussed coding is listed in "2a1.7 Denominator Details" section below.
2a1.8 Denominator Exclusions: ALL MEASURE SPECIFICATION DETAILS REFERENCE THE 2012 PHYSICIAN QUALITY REPORTING SYSTEM MEASURE SPECIFICATION. A patient is identified as a Denominator Exclusions (B) and excluded from the Total Denominator Population (TDP) in the Performance Denominator (PD) calculation if one or more of the following reason (s) exist: There is documentation in the medical record that the patient is over or under weight and is being managed by another provider If the patient has a terminal illness-life expectancy is 6 months or less If the patient is pregnant If the patient refuses BMI measurement If there is any other reason documented in the medical record by the provider explaining why BMI measurement was not appropriate Patient is in an urgent or emergent medical situation where time is of the essence and to delay treatment would jeopardize the

patient's health status.

1.1 Measure Type: Process

2a1. 25-26 Data Source: Administrative claims, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records

2a1.33 Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Population : County or City, Population : National, Population : Regional, Population : State

1.2-1.4 Is this measure paired with another measure? No

De.3 If included in a composite, please identify the composite measure (title and NQF number if endorsed):
n/a

STAFF NOTES (issues or questions regarding any criteria)

Comments on Conditions for Consideration:

Is the measure untested? Yes No If untested, explain how it meets criteria for consideration for time-limited endorsement:

1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (check De.5):

5. Similar/related [endorsed](#) or submitted measures (check 5.1):

Other Criteria:

Staff Reviewer Name(s):

1. IMPACT, OPPORTUNITY, EVIDENCE - IMPORTANCE TO MEASURE AND REPORT

Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement. All three subcriteria must be met to pass this criterion. See [guidance on evidence](#).

Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)

1a. High Impact: H M L I

(The measure directly addresses a specific national health goal/priority identified by DHHS or NPP, or some other high impact aspect of healthcare.)

De.4 Subject/Topic Areas (Check all the areas that apply): Prevention, Prevention : Development/Wellness, Prevention : Obesity, Prevention : Screening

De.5 Cross Cutting Areas (Check all the areas that apply): Population Health, Prevention, Prevention : Obesity, Prevention : Screening

1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers, A leading cause of morbidity/mortality, Frequently performed procedure, High resource use, Patient/societal consequences of poor quality

1a.2 If "Other," please describe:

1a.3 Summary of Evidence of High Impact (Provide epidemiologic or resource use data):

BMI ABOVE NORMAL PARAMETERS

"In 2009, no state met the healthy people 2012 obesity target of 15 percent, and the self reported overall prevalence of obesity among U.S. adults had increased 1.1 percentage points from 2007. Overall self -reported obesity prevalence in the U.S. was 26.7 percent" (CDC, 2010).

Obesity continues to be a public health concern in the United States and throughout the world. In the United States, obesity prevalence doubled among adults between 1980 and 2004 (Flegal, et al., 2002; Ogden, et al, 2006). Obesity is associated with increased risk of a number of conditions, including diabetes mellitus, cardiovascular disease, hypertension, and certain cancers,

and with increased risk of disability and a modestly elevated risk of all-cause mortality. "Obesity is associated with an increased risk of death, particularly in adults younger than age 65 years. Obesity has been shown to reduce life expectancy by 6 to 20 years depending on age and race. Ischemic heart disease, diabetes, cancer (especially liver, kidney, breast, endometrial, prostate and colon), and respiratory diseases are the leading causes of death in persons who are obese"(AHRQ, 2011).

BMI BELOW NORMAL PARAMETERS

Results from the 2009-2010 National Health and Nutrition Examination Survey (NHANES) indicate that an estimated 35.7 percent of adults are obese (CDC, 2012). Although the prevalence of adults in the U.S. who are obese is still high with about one-third of adults obese in 2007-2008, data suggest that the rate of increase for obesity in the U.S. in recent decades may be slowing (Flegal, et al., 2010).

Huffman (2002) states eElderly patients with unintentional weight loss are at higher risk for infection, depression and death. The leading causes of involuntary weight loss are depression (especially in residents of long-term care facilities), cancer (lung and gastrointestinal malignancies), cardiac disorders and benign gastrointestinal diseases. Medications that may cause nausea and vomiting, dysphagia, dysgeusia and anorexia have been implicated. Polypharmacy can cause unintended weight loss, as can psychotropic medication reduction (e.g., by unmasking problems such as anxiety).

1a.4 Citations for Evidence of High Impact cited in 1a.3:

Centers for Disease Control and Prevention (2010). Prevalence of Underweight Among Adults Aged 20 Years and Over: United States, 2007-2008. Retrieved from http://www.cdc.gov/nchs/data/hestat/underweight_adult_07_08/underweight_adult_07_08.pdf

Flegal, K.M., Carroll, M.D., Ogden, C.L., Johnson, C.L. (2002). Prevalence and trends in obesity among US adults, 1999-2000. *Journal of the American Medical Association*, 288: 1723-7

Ogden, C.L., Carroll, M.D., Curtin, L.R., McDowell, M.A., Tabak, C.J., Flegal, K.M. (2006). Prevalence of overweight and obesity in the United States, 1999-2004. *Journal of the American Medical Association*, 295(13): 1549-1555

Agency for Healthcare Research and Quality (2011). Screening for and Management of Obesity and Overweight in Adults. Evidence Synthesis Number 89. Retrieved from <http://www.uspreventiveservicestaskforce.org/uspstf11/obeseadult/obesees.pdf>

Centers for Disease Control and Prevention, National Center for Health Statistics (2012). Prevalence of Obesity in the United States, 2009-2010. NCHS Data Brief, No. 82. Retrieved from <http://www.cdc.gov/nchs/data/databriefs/db82.pdf>

Flegal, K.M., Graubard, B.L., Williamson, D.F., Mitchell, H. G. (2010). Excess Deaths Associated With Underweight, Overweight, and Obesity. *Journal of the American Medical Association*, 293(15): 1861-1867

Huffman, G.B. (2002). Evaluating and Treating Unintentional Weight Loss in the Elderly. *American Family Physician*. 65(4). Retrieved from <http://www.aafp.org/afp/2002/0215/p640.pdf>

1b. Opportunity for Improvement: H M L I

(There is a demonstrated performance gap - variability or overall less than optimal performance)

1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:

Recent literature indicates nearly 50 percent of primary care physician visits did not include a record of the height and weight data necessary to calculate BMI (Ma, et al, 2009).

BMI ABOVE NORMAL PARAMETERS

For clinically obese patients (BMI = 30), 70 percent did not receive a diagnosis of obesity and 63 percent did not receive counseling from their physician (Ma, et al, 2009). Lack of provider documentation of obesity is linked to the absence of counseling patients about weight loss and the health risks of obesity (Waring, et al, 2009)

Although obesity disproportionately affects minorities and the socioeconomically disadvantaged (Ogden, et al, 2006), prior research has shown that clinician diagnosis and treatment of obesity is not consistent with underlying population prevalence Smedley, et al.,

(2002), reported in Bleich, et al. (2010), very low rates of obesity claims among an insured, obese population, particularly for members who were morbidly obese or living in neighborhoods with a higher proportion of Black residents. These findings indicate the need for better systems or incentive structures to encourage more appropriate diagnosis of 11 obese patients in claims data.

Ma, et al (2009) performed a retrospective, cross-sectional analysis of ambulatory visits in the National Ambulatory Medical Care Survey from 2005 and 2006. The study findings on obesity and office-based quality of care concluded the evidence is compelling that obesity is underappreciated in office-based physician practices across the United States (Ma, et al, 2009). Many opportunities are missed for obesity screening and diagnosis, as well as for the prevention and treatment of obesity

BMI BELOW NORMAL PARAMETERS

Ranhoff, et al., (2005) identified using a BMI < 23, resulted in a positive screen for malnutrition (sensitivity 0.86, specificity 0.71), giving 0.75 correctly classified subjects, thus leading to the recommendation that a score of BMI < 23 should be followed by MNA-SF when the aim is to identify poor nutritional status in elderly.

1b.2 Summary of Data Demonstrating Performance Gap (*Variation or overall less than optimal performance across providers*):
[For Maintenance – Descriptive statistics for performance results for this measure - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.]

The description of the claims data for each 6 month time period are as follows:

Performance measure scores demonstrated needed improvement among eligible professionals as the aggregate performance rate based on claims reporting decline. This decline was noted in consecutive reporting periods from 66.1% to 54.3% with increasing numbers of NPIs reporting (1,468 and 3,436, respectively).

Dates of service from 7/1/2008 to 12/31/2008

Aggregate measure performance rate: 49,195/74,445 (66.1%)

Distribution of provider scores (by NPI): N=1,468, Mean = 64.1%, Median=84.3%, SD=40 Range=100

10th percentile: 0%, 25th percentile: 26.9%; 50th percentile: 84.3%; 75th percentile 100.0%

Total Claims Submitted with any G code (G8420, G8417, G8418, G8422, G8421, G8419):117,317

Valid Denominator Criteria: 77,397 (66.0% of total)

Performance Exclusion: 2,952 (3.8% of valid submissions)

Dates of service from 1/1/2009 to 6/30/2009

Aggregate measure performance Rate: 110,701/203,916 (54.3%)

Distribution of provider scores (by NPI): N=3,436, Mean = 54.5%, Median=50.7%, SD=40 Range=100

10th percentile: 0.0%, 25th percentile: 14.6%; 50th percentile: 50.7%; 75th percentile 100.0%

Total Claims Submitted with any G code (G8420, G8417, G8418, G8422, G8421, G8419): 254,827

Valid Denominator Criteria: 209,244 (82.1% of total)

Performance Exclusion: 5,328 (2.6% of valid submissions)

Total tested claims sampled and reviewed: 307 records from 78 providers

Valid denominator criteria: 305/307 (99.3% of total)

Sample Performance Exclusion (claims based): 28 (9.2% of valid)

Measure performance rate (claims based): 59.2%

1b.3 Citations for Data on Performance Gap: **[For Maintenance – Description of the data or sample for measure results reported in 1b.2 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included]**

Ma, J., Xiao, L., & Stafford, R.S. (2009). Adult Obesity and Office-Based of Care in the United States. *Obesity*, 17(5): 1077-1085

Waring, M.E., Roberts, M.B., Parker, D.R., & Eaton, C.B. (2009). Documentation and Management of Overweight and Obesity in Primary Care. *The Journal of the American Board of Family Medicine*, 22 (5): 544-552,

Ogden, C.L., Carroll, M.D., Curtin, L.R., McDowell, M.A., Tabak, C.J., Flegal, K.M. (2006). Prevalence of overweight and obesity in the United States, 1999-2004. *Journal of the American Medical Association*, 295(13): 1549-1555

Bleich, S.N., Clark, J.M., Good, S.M., Huizinga, M.M., & Weiner, J.P. (2010). Variation in Provider Identification of Obesity by Individual-and Neighborhood-Level Characteristics among an Insured Population. *Journal of Obesity*. doi: 10.1155/2010/637829

Ranhoff, A.H., Gjoen, A.U., Mowe, M. (2005). Screening for Malnutrition in Elderly Acute Medical Patients: The Usefulness of MNA-SF. *The Journal of Nutrition, Health & Aging*. 9(4): 221-225

1b.4 Summary of Data on Disparities by Population Group: [For Maintenance –Descriptive statistics for performance results for this measure by population group]

Data analysis can produce provider level performance rates as well as aggregate rates based on any classification and demographic data that can be linked to the provider or patient related to: Race, Gender, Age, Rural/Urban, Underserved/Non-Underserved, and Region. Disparities in performance may be identified by examining these aggregate performance rates.

Aggregate performance rates for the following categories were observed for PQRS claims reporting from 1/1/2009 to 6/30/2009 consisting of 203,916 claims with valid denominator criteria and no performance exclusion. Performance rates represent only those providers who voluntarily reported this measure and cannot be generalized to the population of eligible providers. Disparities data will be displayed as: Disparities category: Performance Rate (sample size)

Rural: 48.8% (n=29,081) Urban: 55.28% (n=174,831)

Urban providers reported more often than rural providers and had a higher aggregate performance rate.

Female: 54.7% (n=117,621) Male: 53.8% (n=86,295)

Medicare claims reporting the measure were predominately female beneficiaries.

Underserved (racial/ethnic minority): 47.4% (n=18,188) Non-underserved: 54.9% (n=184,083) (missing=1645)

Racial and ethnic minority beneficiaries had a higher aggregate performance rate than white beneficiaries.

Race

Asian: 62.3% (n=1680) Black: 43.1% (n=14,555) Hispanic: 70.9% (n=1538) Native American/Pacific Islander: 48.7% (n=415)

White: 54.9% (n=184,083) Other/Unknown: 66.8% (n=1645)

Age Groups

Under 50: 37.4% (n=7749) 50-64: 42.6% (n=16,392) 65-69: 54.3% (n=35,952) 70-74: 55.6% (n=41,171) 75+: 56.9% (n=102,652)

Beneficiaries aged 75 years and older made up more than half of reported claims.

Performance by CMS Region

Providers from CMS Region IV consisting of Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina and Tennessee reported the measure most frequently (n=99,887). Region V was the next highest reporting area consisting of Illinois, Indiana, Michigan, Minnesota, Ohio and Wisconsin (n=29,676). The aggregate performance rate of Region IV providers was 52.6% and for Region V was 46.5%.

[Beta Testing Results: Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up SEE ATTACHEMENT SECTION IV. Analysis of Claims Data).

1b.5 Citations for Data on Disparities Cited in 1b.4: [For Maintenance – Description of the data or sample for measure results reported in 1b.4 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included]

Included with sections 1b.2. Summary of Data Demonstrating Performance Gap & 1b.3 Citations for Data on Performance Gap.

1c. Evidence (Measure focus is a health outcome OR meets the criteria for quantity, quality, consistency of the body of evidence.)

Is the measure focus a health outcome? Yes No If not a health outcome, rate the body of evidence.

Quantity: H M L I Quality: H M L I Consistency: H M L I

Quantity	Quality	Consistency	Does the measure pass subcriterion1c?
M-H	M-H	M-H	Yes <input type="checkbox"/>
L	M-H	M	Yes <input type="checkbox"/> IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No <input type="checkbox"/>
M-H	L	M-H	Yes <input type="checkbox"/> IF potential benefits to patients clearly outweigh potential harms: otherwise No <input type="checkbox"/>
L-M-H	L-M-H	L	No <input type="checkbox"/>
Health outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service			Does the measure pass subcriterion1c? Yes <input type="checkbox"/> IF rationale supports relationship
<p>1c.1 Structure-Process-Outcome Relationship (Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome): process</p> <p>1c.2-3 Type of Evidence (Check all that apply): Clinical Practice Guideline, Selected individual studies (rather than entire body of evidence), Systematic review of body of evidence (other than within guideline development)</p> <p>1c.4 Directness of Evidence to the Specified Measure (State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population): Evidence supports a multi-disciplinary approach to body mass index (BMI) assessment & recommended follow-up based on BMI calculation. Studies explored interventions implemented with outpatient facilities/practices to target negative outcomes of out of normal parameters BMI findings.</p> <p>1c.5 Quantity of Studies in the Body of Evidence (Total number of studies, not articles): Agency for Healthcare Research and Quality. The Guide to Clinical Preventive Services 2010-2011: Recommendations of the U.S. Preventive Services Task Force. Retrieved from http://www.ahrq.gov/clinic/pocketgd1011/pocketgd1011.pdf USPSTF Grade: B Recommendation</p> <p>Agency for Healthcare Research and Quality (2011). Screening for and Management of Obesity and Overweight in Adults. Evidence Synthesis Number 89. Retrieved from http://www.uspreventiveservicestaskforce.org/uspstf11/obeseadult/obesees.pdf SORT Study quality level 1 (good-quality patient-oriented evidence)</p> <p>Bleich, S.N., Clark, J.M., Good, S.M., Huizinga, M.M., & Weiner, J.P. (2010). Variation in Provider Identification of Obesity by Individual-and Neighborhood-Level Characteristics among an Insured Population. Journal of Obesity. doi: 10.1155/2010/637829 Study quality level 2 (limited-quality patient-oriented evidence)</p> <p>Cawley, J., Meyerhoefer, C. (2012). The medical care costs of obesity: An instrumental variables approach. Journal of Health Economics, 31: 219-230 Study quality level 2 (limited-quality patient-oriented evidence)</p> <p>Centers for Disease Control and Prevention (2010). Vital Signs: State-Specific Obesity Prevalence Among Adults – United States, 2009, State-specific prevalence of obesity among adults – United States, 2009. Morbidity and Mortality Weekly Review, Retrieved from http://www.cdc.gov/mmwr/pdf/wk/mm59e0803.pdf Study quality level 2 (limited-quality patient-oriented evidence)</p> <p>Centers for Disease Control and Prevention (2010). Prevalence of Underweight Among Adults Aged 20 Years and Over: United States, 2007-2008. Retrieved from http://www.cdc.gov/nchs/data/hestat/underweight_adult_07_08/underweight_adult_07_08.pdf Study quality level 2 (limited-quality patient-oriented evidence)</p>			

Centers for Disease Control and Prevention, National Center for Health Statistics (2012). Prevalence of Obesity in the United States, 2009-2010. NCHS Data Brief, No. 82. Retrieved from <http://www.cdc.gov/nchs/data/databriefs/db82.pdf>
Study quality level 2 (limited-quality patient-oriented evidence)

Finkelstein, E.A., Trogon, J.G., Cohen, J.W., & Dietz, W. (2009). Annual Medical Spending Attributable To Obesity: Payer-And Service-Specific Estimates. *Health Affairs*, 28(5), w822-w831. doi: 10.1377/hlthaff.28.5.w822
Study quality level 2 (limited-quality patient-oriented evidence)

Flegal, K.M., Carroll, M.D., Ogden, C.L., Johnson, C.L. (2002). Prevalence and trends in obesity among US adults, 1999-2000. *Journal of the American Medical Association*, 288: 1723-7
Study quality level 2 (limited-quality patient-oriented evidence)

Flegal, K.M., Graubard, B.L., Williamson, D.F., Mitchell, H. G. (2010). Excess Deaths Associated With Underweight, Overweight, and Obesity. *Journal of the American Medical Association*, 293(15): 1861-1867
Study quality level 2 (limited-quality patient-oriented evidence)

Huffman, G.B. (2002). Evaluating and Treating Unintentional Weight Loss in the Elderly. *American Family Physician*. 65(4). Retrieved from <http://www.aafp.org/afp/2002/0215/p640.pdf>
Study quality level 3 (other evidence: guideline)

Institute for Clinical Systems Improvement (2011). Health Care Guideline: Prevention and Management of Obesity (Mature Adolescents and Adults). Fifth Edition. Retrieved From http://www.icsi.org/obesity/obesity_3398.html
Study quality level 3 (other evidence: guideline)

Ma, J., Xiao, L., & Stafford, R.S. (2009). Adult Obesity and Office-Based of Care in the United States. *Obesity*, 17(5): 1077-1085.
Study quality level 2 (limited-quality patient-oriented evidence)

Ogden, C.L., Carroll, M.D., Curtin, L.R., McDowell, M.A., Tabak, C.J., Flegal, K.M. (2006). Prevalence of overweight and obesity in the United States, 1999-2004. *Journal of the American Medical Association*, 295(13): 1549-1555.
Study quality level 2 (limited-quality patient-oriented evidence)

Ranhoff, A.H., Gjoen, A.U., Mowe, M. (2005). Screening for Malnutrition in Elderly Acute Medical Patients: The Usefulness of MNA-SF. *The Journal of Nutrition, Health & Aging*. 9(4): 221-225.
Study quality level 2 (limited-quality patient-oriented evidence)

Tsai, A.G., Williamson, D.F., & Glick, H.A. (2010). Direct medical cost of overweight and obesity in the USA: a quantitative systematic review. Retrieved from <http://www3.interscience.wiley.com/journal/123233768/abstract?CRETRY=1&SRETRY=0>
SORT Study quality level 1 (good-quality patient-oriented evidence)

Waring, M.E., Roberts, M.B., Parker, D.R., & Eaton, C.B. (2009). Documentation and Management of Overweight and Obesity in Primary Care. *The Journal of the American Board of Family Medicine*, 22 (5): 544-552.
Study quality level 2 (limited-quality patient-oriented evidence)

1c.6 Quality of Body of Evidence (*Summarize the certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence resulting from study factors. Please address: a) study design/flaws; b) directness/indirectness of the evidence to this measure (e.g., interventions, comparisons, outcomes assessed, population included in the evidence); and c) imprecision/wide confidence intervals due to few patients or events*):

The body of evidence consists of 17 studies. Two studies have SORT Study quality level 1: good-quality patient-oriented evidence (AHRQ, 2011 & Tsai et al., 2010), 1 study has a USPSTF Grade: B Recommendation (AHRQ [Guideline], 2011), 12 studies have SORT Study quality level 2: limited-quality patient-oriented evidence (Bleich et al., 2010; Cawley, 2012; CDC, 2010; CDC, 2010; CDC, 2010; Finkelstein et al., 2009; Flegal et al., 2002; Flegal, 2010; Ma et al., 2009; Ogden et al., 2006; Ranhoff et al., 2005; &

Waring et al., 2009), and 2 have Study quality level 3 (other evidence: guideline). The evidence bears directly on the importance, benchmarking, performance gaps and disparities of BMI calculation and interventions in the outpatient setting and the potential reduction of negative outcomes with declines in obesity and health improvements for underweight populations. Since the studies show consistently statistically significant effects, there are no issues of "imprecision/wide confidence intervals due to few patients or events".

1c.7 Consistency of Results across Studies (Summarize the consistency of the magnitude and direction of the effect):

Consistency of results across studies: While the magnitude of the effects varies from study to study, the effects are consistently positive.

1c.8 Net Benefit (Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms):

Studies show consistent benefits while detecting no harm and yielding consistent net benefits. Any improvement in improved BMI calculation and appropriate follow up net benefit to patients.

1c.9 Grading of Strength/Quality of the Body of Evidence. Has the body of evidence been graded? **Yes**

1c.10 If body of evidence graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias:

Albert G. Crawford, PhD, MBA, MSIS Associate Professor
Jefferson School of Population Health
1015 Walnut Street, Suite 115
Philadelphia, PA 19107

Not disclosures or bias to report

1c.11 System Used for Grading the Body of Evidence: Other

1c.12 If other, identify and describe the grading scale with definitions:

The Strength of Recommendation Taxonomy (SORT)

An A-level recommendation is based on consistent and good-quality patient-oriented evidence; a B-level recommendation is based on inconsistent or limited-quality patient-oriented evidence; and a C-level recommendation is based on consensus, usual practice, opinion, disease oriented evidence, or case series for studies of diagnosis, treatment, prevention, or screening. The quality of individual studies is rated 1, 2, or 3; numbers are used to distinguish ratings of individual studies from the letters A, B, and C used to evaluate the strength of a recommendation based on a body of evidence.

1c.13 Grade Assigned to the Body of Evidence:

Overall Grading: SORT Strength of Recommendation A: consistent, good-quality patient-oriented evidence. Albert G. Crawford, PhD, MBA, MSIS

1c.14 Summary of Controversy/Contradictory Evidence:

Environmental scan and empirical review did not reveal any relevant controversial or contradictory evidence.

1c.15 Citations for Evidence other than Guidelines(Guidelines addressed below):

N/A

1c.16 Quote verbatim, the specific guideline recommendation (Including guideline # and/or page #):

Although multiple clinical recommendations addressing obesity have been developed by professional organizations, societies and

associations, two recommendations, which exemplify the intent of the measure and address the numerator and denominator, have been identified.

The US Preventive Health Services Task Force (USPSTF) The Guide to Clinical Preventive Services, 2010-2011 recommends that clinicians screen all adult patients for obesity and offer intensive counseling and behavioral interventions to promote sustained weight loss for obese adults (Level Evidence B).

Institute for Clinical Systems Improvement (ICSI, 2011 Prevention and Management of Obesity (Mature Adolescents and Adults) provides the following guidance:

- Calculate the body mass index; classify the individual based on the body mass index categories. Educate patients about their body mass index and their associated risks.
- Weight management requires a team approach. Be aware of clinical and community resources. The patient needs to have an ongoing therapeutic relationship and follow-up with a health care team.
- Weight control is a lifelong commitment, and the health care team can assist with setting specific goals with the patient

1c.17 Clinical Practice Guideline Citation: Agency for Healthcare Research and Quality. The Guide to Clinical Preventive Services 2010-2011: Recommendations of the U.S. Preventive Services Task Force. Retrieved from <http://www.ahrq.gov/clinic/pocketgd1011/pocketgd1011.pdf>

Institute for Clinical Systems Improvement (2011). Health Care Guideline: Prevention and Management of Obesity (Mature Adolescents and Adults). Fifth Edition. Retrieved From http://www.icsi.org/obesity/obesity_3398.html

1c.18 National Guideline Clearinghouse or other URL: http://www.icsi.org/obesity/obesity_3398.html
<http://www.uspreventiveservicestaskforce.org/3rduspstf/obesity/obesrr.htm>

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? **Yes**

1c.20 If guideline recommendation graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias:

Albert G. Crawford, PhD, MBA, MSIS Associate Professor - no disclosures or bias to report

1c.21 System Used for Grading the Strength of Guideline Recommendation: USPSTF

1c.22 If other, identify and describe the grading scale with definitions:

1c.23 Grade Assigned to the Recommendation:

SORT Strength A: Consistent, good-quality patient-oriented evidence

1c.24 Rationale for Using this Guideline Over Others:

The US Preventive Health Services Task Force (USPSTF) (2011) and Institute for Clinical Systems Improvement (ICSI) (2011) guidelines are the most up-to-date and also the ones best supported by high-quality research.

Based on the NQF descriptions for rating the evidence, what was the developer's assessment of the quantity, quality, and consistency of the body of evidence?

1c.25 Quantity: **Moderate** **1c.26 Quality:** **Moderate** **1c.27 Consistency:** **Moderate**

1c.28 Attach evidence submission form: **sent to E. Munthali via email due to uploading errors**

1c.29 Attach appendix for supplemental materials: **sent to E. Munthali via email due to uploading errors**

Was the threshold criterion, *Importance to Measure and Report*, met?

(1a & 1b must be rated moderate or high and 1c yes) Yes No

Provide rationale based on specific subcriteria:

**For a new measure if the Committee votes NO, then STOP.
For a measure undergoing endorsement maintenance, if the Committee votes NO because of 1b. (no opportunity for improvement), it may be considered for continued endorsement and all criteria need to be evaluated.**

2. RELIABILITY & VALIDITY - SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. **(evaluation criteria)**

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate field. Supplemental materials may be referenced or attached in item 2.1. See [guidance on measure testing](#).

S.1 Measure Web Page (In the future, NQF will require measure stewards to provide a URL link to a web page where current detailed specifications can be obtained). Do you have a web page where current detailed specifications for this measure can be obtained? Yes

S.2 If yes, provide web page URL:

https://www.cms.gov/apps/ama/license.asp?file=/PQRS/downloads/2012_PhysQualRptg_IndividualClaimsRegistry_Specs_SupportingDocs_01162012.zip

2a. RELIABILITY. Precise Specifications and Reliability Testing: H M L I

2a1. Precise Measure Specifications. (The measure specifications precise and unambiguous.)

2a1.1 Numerator Statement (Brief, narrative description of the measure focus or what is being measured about the target population, e.g., cases from the target population with the target process, condition, event, or outcome):

[ALL MEASURE SPECIFICATION DETAILS REFERENCE THE 2012 PHYSICIAN QUALITY REPORTING SYSTEM MEASURE SPECIFICATION.](#)

Patients with BMI calculated within the past six months or during the current visit and a follow-up plan documented if the BMI is outside of parameters

2a1.2 Numerator Time Window (The time period in which the target process, condition, event, or outcome is eligible for inclusion):

This measure is to be reported a minimum of once per reporting period for patients seen during the reporting period. There is no diagnosis associated with this measure. This measure may be reported by eligible professionals who perform the quality actions described in the measure based on the services provided and the measure-specific denominator coding. BMI measured and documented in the medical record may be reported if done in the provider's office/facility or if BMI calculation within the past six months is documented in outside medical records obtained by the provider. The documentation of a follow up plan should be based on the most recent calculated BMI.

2a1.3 Numerator Details (All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, codes with descriptors, and/or specific data collection items/responses):

For the purposes of calculating performance, the Numerator (A) is defined by providers reporting the clinical quality action was performed. For this measure, performing the clinical quality action is numerator HCPCS G8420, G8417 & G8418. All discussed coding detail is listed in "2a1.7. Denominator Details" section below.

2a1.4 Denominator Statement (Brief, narrative description of the target population being measured):

[ALL MEASURE SPECIFICATION DETAILS REFERENCE THE 2012 PHYSICIAN QUALITY REPORTING SYSTEM MEASURE SPECIFICATION.](#)

All patients aged 18 years and older on date of encounter seen during the 12 month reporting period with one or more denominator CPT or HCPCS encounter codes reported on the Medicare Part B Claims submission for the encounter along with one of the 6 numerator HCPCS clinical quality codes. All discussed coding is listed in "2a1.7 Denominator Details" section below.

2a1.5 Target Population Category (Check all the populations for which the measure is specified and tested if any):
Adult/Elderly Care, Populations at Risk, Senior Care

2a1.6 Denominator Time Window (The time period in which cases are eligible for inclusion):
All patients aged 18 years and older at the time of the encounter seen during the 12 month reporting period.

2a1.7 Denominator Details (All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses):

The Total Denominator Population (TDP) is defined with the following criteria: 1) patient's age at the time of the encounter 2) encounter date within the 12 month reporting period 3) denominator CPT or HCPCS encounter codes AND 4) provider reported HCPCS numerator clinical quality code described below (G8420, G8417, G8418, G8422, G8421 & G8419).

TOTAL DENOMINATOR POPULATION

Patients aged 18 years and older on the date of the encounter

AND

Patient encounters during the 12 month reporting period with the following CPT or HCPCS encounter codes: 90801, 90802, 90804, 90805, 90806, 90807, 90808, 90809, 97001, 97003, 97802, 97803, 98960, 99201, 99202, 99203, 99204, 99205, 99212, 99213, 99214, 99215, D7140, D7210, G0101, G0108, G0270, G0271, G0402, G0438, G0439

AND

Patient encounters with the following HCPCS numerator clinical quality codes: G8420, G8417, G8418, G8422, G8421 & G8419
HCPCS NUMERATOR CLINICAL QUALITY CODES (6)

PERFORMANCE PASS CLINICAL QUALITY CODES (3)

BMI Calculated as Normal, No Follow-Up Plan Required

G8420: Calculated BMI within normal parameters and documented

BMI Calculated Above Upper Normal Parameters, Follow-Up Documented

G8417: Calculated BMI above the upper parameter and a follow-up plan was documented in the medical record

BMI Calculated Below Lower Normal Parameters, Follow-Up Documented

G8418: Calculated BMI below the lower parameter and a follow-up plan was documented in the medical record

DENOMINATOR EXCLUSION (B) CLINICAL QUALITY CODE (1)

BMI not Calculated, Patient not Eligible/not Appropriate

G8422: Patient not eligible for BMI calculation

PERFORMANCE FAILURE CLINICAL QUALITY CODES (2)

BMI not Calculated, Reason not Specified

G8421: BMI not calculated

BMI Calculated Outside Normal Parameters, Follow-Up Plan not Documented, Reason not Specified

G8419: Calculated BMI outside normal parameters, no follow-up plan documented in the medical record

2a1.8 Denominator Exclusions (Brief narrative description of exclusions from the target population):

ALL MEASURE SPECIFICATION DETAILS REFERENCE THE 2012 PHYSICIAN QUALITY REPORTING SYSTEM MEASURE SPECIFICATION.

A patient is identified as a Denominator Exclusions (B) and excluded from the Total Denominator Population (TDP) in the

Performance Denominator (PD) calculation if one or more of the following reason (s) exist:

There is documentation in the medical record that the patient is over or under weight and is being managed by another provider

If the patient has a terminal illness-life expectancy is 6 months or less

If the patient is pregnant

If the patient refuses BMI measurement

If there is any other reason documented in the medical record by the provider explaining why BMI measurement was not appropriate

Patient is in an urgent or emergent medical situation where time is of the essence and to delay treatment would jeopardize the patient's health status.

2a1.9 Denominator Exclusion Details (All information required to identify and calculate exclusions from the denominator such as definitions, codes with descriptors, and/or specific data collection items/responses):

Denominator Exclusions (B) are identified with the following provider reported HCPCS numerator clinical quality code:

BMI not Calculated, Patient not Eligible/not Appropriate

G8422 Patient not eligible for BMI calculation

DENOMINATOR EXCLUSION CALCULATION:

Denominator Exclusions (B)(G8422)/Total Denominator Population (TDP)(G8420, G8417, G8418, G8421, G8419 & G8422)

2a1.10 Stratification Details/Variables (All information required to stratify the measure results including the stratification variables, codes with descriptors, definitions, and/or specific data collection items/responses):

No stratification. All eligible patients are subject to the same numerator criteria.

2a1.11 Risk Adjustment Type (Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13): No risk adjustment or risk stratification **2a1.12 If "Other," please describe:**

2a1.13 Statistical Risk Model and Variables (Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development should be addressed in 2b4.):

n/a

2a1.14-16 Detailed Risk Model Available at Web page URL (or attachment). Include coefficients, equations, codes with descriptors, definitions, and/or specific data collection items/responses. Attach documents only if they are not available on a webpage and keep attached file to 5 MB or less. NQF strongly prefers you make documents available at a Web page URL. Please supply login/password if needed:

URL

n/a

n/a

2a1.17-18. Type of Score: Rate/proportion

2a1.19 Interpretation of Score (Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score): Better quality = Higher score

2a1.20 Calculation Algorithm/Measure Logic(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.):

THIS SECTION PROVIDES DEFINITIONS & FORMULAS FOR THE NUMERATOR (A), TOTAL DENOMINATOR POPULATION (TDP), DENOMINATOR EXCLUSIONS (B) CALCULATION & PERFORMANCE DENOMINATOR (PD) CALCULATION.

NUMERATOR (A): HCPCS Clinical Quality Codes G8420, G8417 & G8418

TOTAL DENOMINATOR POPULATION (TDP): Patient aged 18 years and older on the date of the encounter of the 12-month reporting period, with denominator defined encounter codes & Medicare Part B Claims reported HCPCS Clinical Quality Codes G8420, G8417, G8418, G8422, G8421 & G8419

DENOMINATOR EXCLUSION CALCULATION:

Denominator Exclusion (B): # of patients with valid exclusions

G8422 / # TDP

PERFORMANCE DENOMINATOR CALCULATION:

Performance Denominator (B): Patients meeting criteria for performance denominator calculation

A / (# TDP - # B)

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:

URL

Please see attached "NQF 0421 Endorsement - Quality Insights of Pennsylvania 050112" document on page 46. Attachment error noted.

n/a

2a1.24 Sampling (Survey) Methodology. If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate):

n/a

2a1.25 Data Source (Check all the sources for which the measure is specified and tested). If other, please describe:

Administrative claims, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records

2a1.26 Data Source/Data Collection Instrument (Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.):

Medicare Part B Claims Data is provided for testing purposes. This measure is also EHR retooled. Per NQF permission, the feasibility, reliability & validity testing results will be provided with the 2013 annual measure update.

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment:

URL

Please see attached "PQRS_128_NQF_0421_PartB_claims_AdHocRecordLayout" document on page 44 of "NQF_0421_Endorsement_Quality_Insights_of_Pennsylvania.pdf". Attachment error noted.

n/a

2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:

URL

Please see attached "2012 Specification Coding" AND "2009 Specification Coding" on pages 27-41 of "NQF_0421_Endorsement_Quality_Insights_of_Pennsylvania.pdf". Attachment error noted.

n/a

2a1.33 Level of Analysis (Check the levels of analysis for which the measure is specified and tested): Clinician : Group/Practice,

Clinician : Individual, Population : County or City, Population : National, Population : Regional, Population : State

2a1.34-35 Care Setting (Check all the settings for which the measure is specified and tested): Ambulatory Care : Clinician Office/Clinic, Ambulatory Care : Outpatient Rehabilitation, Behavioral Health/Psychiatric : Outpatient, Home Health, Other: Dental & Domiciliary Care

2a2. Reliability Testing. (Reliability testing was conducted with appropriate method, scope, and adequate demonstration of reliability.)

2a2.1 Data/Sample (Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

Time period: 1/1/2009 – 6/30/2009

Claim Type: Claim Carrier (B)

Criteria: Any HCPCS Line code in the following string: G8420, G8417, G8418, G8422, G8421, G8419

Additional fields requested to the standard layout: LINE_PRCSG_IND (included in the detail file), beneficiary name, beneficiary DOB, beneficiary DOD, beneficiary gender, beneficiary HIC, and beneficiary race.

NPIs with fewer than ten (10) claims were removed from the dataset. A simple random sample of records for approximately 150 NPIs was drawn. From those 150 NPIs, a random sample of approximately 600 claims was identified. The records were then stratified by the business location address listed in the NPI registry so the maximum number of records from each business location was limited to 10 records. This limitation was set so the providers would not see this task as too burdensome and would be more likely to send in their records.

Randomly selected providers were mailed a letter requesting they provide the documentation to support the assignment of the numerator/G code submitted on the claim. The first request for data was mailed to the selected providers on March 9, 2010. A subsequent reminder letter for those providers who had not mailed their documentation was sent on April 16, 2010

Data Sample Response Rates:

Number of records requested / returned / reviewed: 603/309/307 Provider response rate 51.2%

Number of provider requested / returned / reviewed: 154/89/78 Provider response rate 57.8%

2a2.2 Analytic Method (Describe method of reliability testing & rationale):

Crude agreement rates were calculated along with prevalence adjusted kappa (PAK), Cohen's kappa values and corresponding confidence intervals. Cohen's kappa represents chance-corrected proportional agreement. High prevalence of responses in a small number of cells is known to produce unexpected results known as the "kappa paradox" When the prevalence of a rating in the population is very high or low, which was noted in the testing of this measure, the value of kappa may indicate poor reliability even with a high observed proportion of agreement. In such cases, as with this measure, PAK is shown to provide an additional interpretation of agreement when the prevalence of responses is concentrated in a small number of cells.

Landis and Koch (1977) have proposed the following as standards for strength of agreement for the kappa coefficient: [less than or equal to] 0=poor, .01-.20=slight, .21-.40=fair, .41-.60=moderate, .61-.80=substantial and .81-1 =almost perfect (high). These categories are informal.

2a2.3 Testing Results (Reliability statistics, assessment of adequacy in the context of norms for the test conducted):

Overall Reliability:

Numerator: 76.1% agreement, PAK=.54 (.45 - .63), Kappa=.54 (.45 - .63)

Denominator Exclusions: 93.4% agreement, PAK=.87 (.81 - .92)

Kappa .45 (.25 - .64), Valid Denominator Criteria: 305 / 307 99.3%

Inter-Rater Reliability:

Numerator:91.8% agreement, PAK=.84 (.68-.99), Kappa=.84(.68-.99)

Denominator Exclusions: 98.0% agreement, PAK=.96 (.88-1.00), Kappa .00 (.00-.00) Valid Denominator Criteria: 50/50 (100%)

All records without valid denominator criteria were removed prior to reliability assessment. Denominator agreement was 100%.

Reporting of this measure demonstrates moderate reliability and there is substantial IRR agreement between ALPS and Quality Insights in the testing of this measure. Further analysis of reported claims discrepancies demonstrate provider education on the documentation requirements for this measure may improve reporting reliability including enhancing language to the measure statements to stress that the BMI and follow-up, if applicable. With respect to claims data analysis, additional education may be warranted to further clarify for providers who are eligible to report the measure based on the comprehensive denominator eligibility criteria. [Beta Testing Results : Preventive Care and Screening : Body Mass Index (BMI) Screening and Follow Up SEE ATTACHMENT SECTION II. Reliability Testing].

2b. VALIDITY. Validity, Testing, including all Threats to Validity: H M L I

2b1.1 Describe how the measure specifications (measure focus, target population, and exclusions) are consistent with the evidence cited in support of the measure focus (criterion 1c) and identify any differences from the evidence:

Quality Insights of Pennsylvania conducts an Environmental Scan to evaluate the most current research and evidence-based guidelines. The TEP, composed of subject matter specialists and experts with technical measure expertise evaluates the results of the review and provides recommendations based on the scientific merits of the evidence using the Strength of Recommendation Taxonomy (SORT). The TEP also reviews and establishes the measure's ability to capture what it is designed to capture using a consensus process.

The initial measure development process included alpha-testing in the field with select providers and a public comment period. During the Reliability Testing, Quality Insights again convened a TEP for Environmental Scan review as well as a detailed analysis of beta testing results. Based on the process of multiple stakeholder input, expert panel discussion and public comment, face and content validity of CMS/Quality Insights measures can be assumed to be established.

2b2. Validity Testing. (Validity testing was conducted with appropriate method, scope, and adequate demonstration of validity.)

2b2.1 Data/Sample (Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

See 2b1.1

2b2.2 Analytic Method (Describe method of validity testing and rationale; if face validity, describe systematic assessment):

See 2b1.1

2b2.3 Testing Results (Statistical results, assessment of adequacy in the context of norms for the test conducted; if face validity, describe results of systematic assessment):

See 2b1.1

POTENTIAL THREATS TO VALIDITY. (All potential threats to validity were appropriately tested with adequate results.)

2b3. Measure Exclusions. (Exclusions were supported by the clinical evidence in 1c or appropriately tested with results demonstrating the need to specify them.)

2b3.1 Data/Sample for analysis of exclusions (Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

Claims data from 7/1/2008 – 6/30/2009. Testing performed on sample (See 2a2.3 - Testing Results).

2b3.2 Analytic Method (Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference):

Claims data were analyzed for frequency of reported exclusions and impact on performance scores.

Reliability of exception code assignment was assessed (See 2a2.3 - Testing Results). Crude agreement rates were calculated along with kappa values and corresponding confidence intervals. [Beta Testing Results : Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow Up: SEE ATTACHMENT SECTION II. Reliability Testing].

2b3.3 Results (Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):

Overall reliability Performance Exclusions: There were 305 cases in the testing sample with valid denominator criteria. Based on

codes submitted with claims data there were 28 (9.2%) denominator exclusions. Agreement was assessed as follows:

Overall Reliability:

Performance Exclusions: 93.4% agreement, PAK=.87 (.81 - .92) Kappa .45 (.25 - .64), Inter-Rater Reliability:

Performance Exclusions: 98.0% agreement, PAK=.96 (.88 – 1.00) Kappa .00 (.00-.00)

2b4. Risk Adjustment Strategy. (For outcome measures, adjustment for differences in case mix (severity) across measured entities was appropriately tested with adequate results.)

2b4.1 Data/Sample (Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

N/A

2b4.2 Analytic Method (Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables):

N/A

2b4.3 Testing Results (Statistical risk model: Provide quantitative assessment of relative contribution of model risk factors; risk model performance metrics including cross-validation discrimination and calibration statistics, calibration curve and risk decile plot, and assessment of adequacy in the context of norms for risk models. Risk stratification: Provide quantitative assessment of relationship of risk factors to the outcome and differences in outcomes among the strata):

N/A

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment: The processes being reported in this measure would not be influenced by patient characteristics, setting or other factors outside of the provider's control.

2b5. Identification of Meaningful Differences in Performance. (The performance measure scores were appropriately analyzed and discriminated meaningful differences in quality.)

2b5.1 Data/Sample (Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

The description of the claims data for each 6 month time period are as follows:

Dates of service from 7/1/2008 to 12/31/2008

Total Claims Submitted with any G code (G8420, G8417, G8418, G8422, G8421, G8419):117,317

Valid Denominator Criteria: 77,397 (66.0% of total)

Performance Exclusion: 2,952 (3.8% of valid submissions)

Dates of service from 1/1/2009 to 6/30/2009

Total Claims Submitted with any G code (G8420, G8417, G8418, G8422, G8421, G8419): 254,827

Valid Denominator Criteria: 209,244 (82.1% of total)

Performance Exclusion: 5,328 (2.6% of valid submissions)

Total claims sampled and reviewed: 307 records from 78 providers

Valid denominator criteria: 305/307 (99.3% of total)

Sample Performance Exclusion (claims based): 28 (9.2% of valid)

Measure performance rate (claims based): 59.2%

2b5.2 Analytic Method (Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance):

Aggregate and provider (NPI) performance rates were calculated from Part B claims with dates of service for two consecutive six month periods. Data from the testing sample were not analyzed at the provider level. Performance rates are derived from G codes submitted for the Physician Quality Reporting System (formerly PQRI). Code submissions are voluntary and providers who report may not be representative of all eligible professionals. Performance rates cannot be generalized to the population.

2b5.3 Results (Provide measure performance results/scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningful differences in performance):

Performance measure scores demonstrated needed improvement among eligible providers as the aggregate performance rate based on claims reporting decline. This decline was noted in consecutive reporting periods from 66.1% to 54.3% with increasing numbers of NPIs reporting (1468 and 3436, respectively).

Dates of service from 7/1/2008 to 12/31/2008

Aggregate measure performance rate: 49,195/74,445 (66.1%)

Distribution of provider scores (by NPI): N=1,468, Mean = 64.1%, Median=84.3%, SD=40 Range=100

10th percentile: 0%, 25th percentile: 26.9%; 50th percentile: 84.3%; 75th percentile 100.0%

Dates of service from 1/1/2009 to 6/30/2009

Aggregate measure performance Rate: 110,701/203,916 (54.3%)

Distribution of provider scores (by NPI): N=3,436, Mean = 54.5%, Median=50.7%, SD=40 Range=100

10th percentile: 0.0%, 25th percentile: 14.6%; 50th percentile: 50.7%; 75th percentile 100.0%

Testing sample

Measure performance rate (claims based): 164/277 (59.2%)

2b6. Comparability of Multiple Data Sources/Methods. (If specified for more than one data source, the various approaches result in comparable scores.)

2b6.1 Data/Sample (Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

N/A

2b6.2 Analytic Method (Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure):

N/A

2b6.3 Testing Results (Provide statistical results, e.g., correlation statistics, comparison of rankings; assessment of adequacy in the context of norms for the test conducted):

N/A

2c. Disparities in Care: H M L I NA (If applicable, the measure specifications allow identification of disparities.)

2c.1 If measure is stratified for disparities, provide stratified results (Scores by stratified categories/cohorts): N/A

2c.2 If disparities have been reported/identified (e.g., in 1b), but measure is not specified to detect disparities, please explain:

Data analysis can produce provider level performance rates as well as aggregate rates based on any classification and demographic data that can be linked to the provider or patient related to: Race, Gender, Age, Rural/Urban, Underserved/Non-Underserved, and Region. Disparities in performance may be identified by examining these aggregate performance rates.

Aggregate performance rates for the following categories were observed for PQRS claims reporting from 1/1/2009 to 6/30/2009 consisting of 203,916 claims with valid denominator criteria and no performance exclusion. Performance rates represent only those providers who voluntarily reported this measure and cannot be generalized to the population of eligible providers. Disparities data will be displayed as: Disparities category: Performance Rate (sample size)

Rural: 48.8% (n=29,081) Urban: 55.28% (n=174,831)

Urban providers reported more often than rural providers and had a higher aggregate performance rate.

Female: 54.7% (n=117,621) Male: 53.8% (n=86,295)

Medicare claims reporting the measure were predominately female beneficiaries.

Underserved (racial/ethnic minority): 47.4% (n=18,188) Non-underserved: 54.9% (n=184,083) (missing=1645)
Racial and ethnic minority beneficiaries had a higher aggregate performance rate than white beneficiaries.

Race

Asian: 62.3% (n=1680) Black: 43.1% (n=14,555) Hispanic: 70.9% (n=1538) Native American/Pacific Islander: 48.7% (n=415)
White: 54.9% (n=184,083) Other/Unknown: 66.8% (n=1645)

Age Groups

Under 50: 37.4% (n=7749) 50-64: 42.6% (n=16,392) 65-69: 54.3% (n=35,952) 70-74: 55.6% (n=41,171) 75+: 56.9% (n=102,652)

Beneficiaries aged 75 years and older made up more than half of reported claims.

Performance by CMS Region

Providers from CMS Region IV consisting of Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina and Tennessee reported the measure most frequently (n=99,887). Region V was the next highest reporting area consisting of Illinois, Indiana, Michigan, Minnesota, Ohio and Wisconsin (n=29,676). The aggregate performance rate of Region IV providers was 52.6% and for Region V was 46.5%.

[Beta Testing Results: Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up SEE ATTACHEMENT SECTION IV. Analysis of Claims Data].

2.1-2.3 Supplemental Testing Methodology Information:

URL

Attachment error noted. Emailed "NQF_0421_Endorsement_Quality_Insights_of_Pennsylvania" to E. Munthali
n/a

Steering Committee: Overall, was the criterion, *Scientific Acceptability of Measure Properties*, met? (Reliability and Validity must be rated moderate or high) Yes No

Provide rationale based on specific subcriteria:

If the Committee votes No, STOP

3. USABILITY

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

C.1 Intended Actual/Planned Use (Check all the planned uses for which the measure is intended): Payment Program, Public Health/Disease Surveillance, Public Reporting, Quality Improvement (Internal to the specific organization), Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

3.1 Current Use (Check all that apply; for any that are checked, provide the specific program information in the following questions): Public Reporting, Payment Program, Public Health/ Disease Surveillance, Quality Improvement with Benchmarking (external benchmarking to multiple organizations), Quality Improvement (Internal to the specific organization)

3a. Usefulness for Public Reporting: H M L I

(The measure is meaningful, understandable and useful for public reporting.)

3a.1. Use in Public Reporting - disclosure of performance results to the public at large (If used in a public reporting program, provide name of program(s), locations, Web page URL(s)). If not publicly reported in a national or community program, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: [**For Maintenance** – If not publicly reported, describe progress made toward achieving disclosure of performance results to the public at large and expected date for public reporting; provide rationale why continued endorsement should be considered.]

Physician Quality Reporting System <http://www.cms.gov/PQRS>

EHR Incentive Program <https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/index.html>

Value Based Modifier <http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/sharesavingsprogram/index.html?redirect=/sharesavingsprogram/>

This measure is used in a public reporting program on the CMS Physician Compare website. Individual level provider performance is anticipated for publication in 2013 with 2012 performance data at the link provided below.

<http://www.medicare.gov/find-a-doctor/provider-search.aspx>

3a.2. Provide a rationale for why the measure performance results are meaningful, understandable, and useful for public reporting. If usefulness was demonstrated (e.g., focus group, cognitive testing), describe the data, method, and results: [Please see the attached CMS web links for performance reporting.](#)

3.2 Use for other Accountability Functions (payment, certification, accreditation). If used in a public accountability program, provide name of program(s), locations, Web page URL(s): [Physician Quality Reporting System Incentive Program](#)

3b. Usefulness for Quality Improvement: H M L I

(The measure is meaningful, understandable and useful for quality improvement.)

3b.1. Use in QI. If used in quality improvement program, provide name of program(s), locations, Web page URL(s):

[For Maintenance – If not used for QI, indicate the reasons and describe progress toward using performance results for improvement].

This measure is used in a public reporting program on the CMS Physician Compare website. Individual level provider performance is anticipated for publication in 2013 with 2012 performance data at the link provided below.

<http://www.medicare.gov/find-a-doctor/provider-search.aspx>

3b.2. Provide rationale for why the measure performance results are meaningful, understandable, and useful for quality improvement. If usefulness was demonstrated (e.g., QI initiative), describe the data, method and results:

See [Physician Quality Reporting System Overview section at www.cms.gov/pqrs](#)

Feedback reports are generated and available for provider performance review.

Overall, to what extent was the criterion, Usability, met? H M L I

Provide rationale based on specific subcriteria:

4. FEASIBILITY

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

4a. Data Generated as a Byproduct of Care Processes: H M L I

4a.1-2 How are the data elements needed to compute measure scores generated? *(Check all that apply).*

Data used in the measure are:

[generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition, Coded by someone other than person obtaining original information \(e.g., DRG, ICD-9 codes on claims\)](#)

4b. Electronic Sources: H M L I

4b.1 Are the data elements needed for the measure as specified available electronically *(Elements that are needed to compute measure scores are in defined, computer-readable fields):* [ALL data elements are in a combination of electronic sources](#)

4b.2 If ALL data elements are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources:

4c. Susceptibility to Inaccuracies, Errors, or Unintended Consequences: H M L I

4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:

Reporting of this measure demonstrates moderate reliability and there is substantial IRR agreement between ALPS and Quality Insights in the testing of this measure. Further analysis of reported claims discrepancies demonstrate provider education on the documentation requirements for this measure may improve reporting reliability including enhancing language to the measure statements to stress that the BMI and follow-up, if applicable. With respect to claims data analysis, additional education may be warranted to further clarify for providers who are eligible to report the measure based on the comprehensive denominator eligibility criteria.

4d. Data Collection Strategy/Implementation: H M L I

A.2 Please check if either of the following apply (regarding proprietary measures):

4d.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues (e.g., fees for use of proprietary measures):

Quality Insights obtained data from a total of 372,144 claims that were submitted with one of the measure's numerator G codes for encounters between 7/1/2008 and 6/30/2009. In the first 6 months of 2009 3,436 unique providers submitted claims with valid reporting for the measure.

Retooling of this measure for compatibility with EHRs has been completed and implemented in 2011. EHR Testing to be submitted with 2013 NQF annual endorsement update per NQF guidance.

Overall, to what extent was the criterion, *Feasibility*, met? H M L I

Provide rationale based on specific subcriteria:

OVERALL SUITABILITY FOR ENDORSEMENT

Does the measure meet all the NQF criteria for endorsement? Yes No

Rationale:

If the Committee votes No, STOP.

If the Committee votes Yes, the final recommendation is contingent on comparison to related and competing measures.

5. COMPARISON TO RELATED AND COMPETING MEASURES

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure before a final recommendation is made.

5.1 If there are related measures (either same measure focus or target population) or competing measures (both the same measure focus and same target population), list the NQF # and title of all related and/or competing measures:

0023 : Body Mass Index (BMI) in adults > 18 years of age

0024 : Weight Assessment and Counseling for Nutrition and Physical Activity for Children/Adolescents

0689 : Percent of Residents Who Lose Too Much Weight (Long-Stay)

1349 : Child Overweight or Obesity Status Based on Parental Report of Body-Mass-Index (BMI)

5a. Harmonization

5a.1 If this measure has EITHER the same measure focus OR the same target population as [NQF-endorsed measure\(s\)](#): Are the measure specifications completely harmonized? No

5a.2 If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on

interpretability and data collection burden:

1349 reports BMI for 10-17 years classifying weight in underweight, normal, overweight & obese; 0023 is a BMI reporting-only measure with the same age population and setting as 0421 but looks for a 24 month finding and does not recommend follow up for out of parameter findings; 0024 reports only for the same setting as in 0023 in well-child visits for patients aged 2 through 18 years; 0689 reports weight loss of 5% in 1 month & 10% in 6 months for long term care patients > 100 days length of stay.

5b. Competing Measure(s)

5b.1 If this measure has both the same measure focus and the same target population as NQF-endorsed measure(s): Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible):

Expands a screening measure to include an actionable clinical intervention to improve quality of care. 0421 is an adult measure in the outpatient setting looking for BMI measurement every year with recommended follow up provided based BMI findings outside of normal parameters in the last 6 months. No other measure provides for measurement & intervention. This measure is widely adopted in numerous clinical quality programs and is available in claims, registry and electronic health record versions.

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner): Centers for Medicare and Medicaid Services, 7500 Security Boulevard , Mail Stop S3-01-02, Baltimore, Maryland, 21244-1850

Co.2 Point of Contact: Edward Q., Garcia III, MHS, Health Policy Analyst, MMSNQF@hsag.com, 410-786-6738-

Co.3 Measure Developer if different from Measure Steward: Quality Insights of Pennsylvania, 630 Freedom Business Center, Suite 116, King of Prussia, Pennsylvania, 19406

Co.4 Point of Contact: Sharon, Hibay, RN, DNP, shibay@wvmi.org, 877-346-6180-7814

Co.5 Submitter: Sharon, Hibay, RN, DNP, shibay@wvmi.org, 877-346-6180-7814, Quality Insights of Pennsylvania

Co.6 Additional organizations that sponsored/participated in measure development:
Thomas Jefferson University School of Population Health
ALPS Services Inc.

Co.7 Public Contact: Sharon, Hibay, RN, DNP, shibay@wvmi.org, 877-346-6170-7814, Quality Insights of Pennsylvania

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development

Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

Through a collaborative process, the TEP reviewed the current 2012 measure specifications (description, numerator, denominator, definitions, clinical recommendation, and environmental scan); reviewed and considered the Beta Testing results, analysis, findings and recommendations based on testing. TEP Recommended the following actions: BMI Parameter for 65 and older changed from < 22 to < 23; education of providers supported as recommended; Clinical Recommendations of the USPSTF (2011) and ICSI (2011) accepted as supporting the measure appropriately; add underweight literature citation to the rationale and High Impact sections as measure addresses both underweight and overweight; retain all G codes as currently documented; add referral types, including surgeon, to current specification definitions under follow-up – referral (registered dietitian, etc); retain all exceptions listed in the definition of Not Eligible/Not Appropriate; do not delete future appointment, etc as retained in the E H R specifications; at this time – do not further define Plan of Care/Care Plan, Nutrition Counseling, and Prescribe/Administer Medications.

Christina K. Biesemeier, MS, RD, LDN, FADA
TEP Chairperson
Vanderbilt University Medical Center
Director, Clinical Nutrition Services

Rhea Cohn, PT, DPT
Director Workers Compensation Business Operations National Rehabilitation Hospital

Mirean Coleman, MSW, LICSW, CT
Senior Practice Associate
National Association for Social Workers

Jenifer Osorno Fahey, CNM, MSN, MSPH
Assistant Professor, Department of Obstetrics Gynecology and Reproductive Services
University of Maryland School of Medicine

Karen Grant, MD
Texas Healthcare

Anthony W. Hamm, DC, FACO

Kirk Koyama, MSN, RN, PHN, CNS
Indian Health Services-Chinle Comprehensive Health Care Facility

Elisa Marks, OTR/L, CHT
RehabNetOutpatient Center

Gregory M. Martino, PhD

Kathleen Niedert, PhD, MBA, RD, CSG, LD, FADA, NHA/L
Omega Health Associates

Kevin M. Schuer, PA-C, MSPAS, MPH
Center for Enterprise Quality and Safety & the University of Kentucky College of Health Sciences
Lexington, KY 40536-0200

Jan Towers, PhD, NP-C, CRNP, FAANP, FAAN
Director of Health Policy
American Academy of Nurse Practitioners
Office of Health Policy

Ad.2 If adapted, provide title of original measure, NQF # if endorsed, and measure steward. Briefly describe the reasons for adapting the original measure and any work with the original measure steward: n/a

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.3 Year the measure was first released: 2008

Ad.4 Month and Year of most recent revision: 10/2012

Ad.5 What is your frequency for review/update of this measure? Annually

Ad.6 When is the next scheduled review/update for this measure? 05/2012

Ad.7 Copyright statement: CPT only copyright 2008-2011 American Medical Association. All rights reserved.
CPT is a registered trademark of the American Medical Association.

Applicable FARS/DFARS Apply to Government Use.

Fee schedules, relative value units, conversion factors and/or related components are not assigned by the AMA, are not part of CPT, and the AMA is not recommending their use. The AMA does not directly or indirectly practice medicine or dispense medical services. The AMA assumes no liability for data contained or not contained herein.

Ad.8 Disclaimers: The measure and specification are provided "as is" without warranty of any kind.

NQF #0421 Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up, Last Updated Date: May 08, 2012

Ad.9 Additional Information/Comments: Attachment error noted. Emailed supplementary documents "NQF_0421_Endorsement_Quality_Insights_of_Pennsylvania" AND "NQF_0421_#1.zip" AND "NQF_0421_#2.zip" to E. Munthali.

Date of Submission (MM/DD/YY): 05/08/2012